Bruno Filipe Madeira de Albuquerque

GENETICS AND DYNAMICS OF piRNA INDUCED SILENCING

Tese de Candidatura ao grau de Doutor em Biologia Básica e Aplicada submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – René François Ketting
Categoria – Professor Catedrático
Afiliação – Universidade Johannes Gutenberg

Co-orientador – Maria Alexandra Marques Moreira Mourão do Carmo
Categoria – Investigadora Principal
Afiliação – Universidade do Porto

Para a avó Laura

Contents

Resumo

O objectivo mais básico de qualquer organismo é a transmissão do seu material genético a futuras gerações. Como tal, um dos desafios impostos às células germinais é a manutenção da sua integridade genómica face a parasitas moleculares como transposões e retrovírus, que, se não forem controlados podem causar infertilidade. Para tal, os organismos desenvolveram diversas estratégias para silenciar material genético exógeno, como é o caso do uso de proteínas Argonautas da classe das PIWI ("P-element Induced Wimpy testis") e os pequenos RNAs a que estas se associam (piRNAs). Este mecanismo é designado por via dos piRNAs ou via dos PIWI. No nematode *C. elegans*, PRG-1 ('PIWI Related Gene-1), uma proteína ortóloga de PIWI, utiliza um leque de mais de 30000 piRNAs para identificar por complementaridade sequências externas. Após identificar um transcrito alvo, PRG-1 recruta proteínas do complexo "mutator" que estabelecem o silenciamento do transcrito alvo. Este silenciamento é caracterizado pela síntese de uma segunda classe de pequenos RNAs, que são complementares à sequência do transcrito na proximidade do local de reconhecimento por PRG-1. Estes pequenos RNAs são apelidados de "22G", pois possuem 22 nucleótidos de comprimento e contêm uma guanina na posição 5'. Este mecanismo é chamado de via secundária e os 22Gs classificados como pequenos RNAs secundários. Através de um mecanismo ainda desconhecido, a via secundária é capaz de estabelecer um silenciamento nuclear que se torna independente de PRG-1. Este estado, chamado de RNAe ("RNA induced epigenetic silencing"), é hereditário e dependente dos genes mutator, de pequenos RNAs 22G de outra proteína Argonauta, HRDE-1 (Heritable RNAi Deficient-1).

O uso de mais de 30000 "piRNAs" diferentes que identificam o alvo por complementaridade imperfeita possibilita o reconhecimento de virtualmente qualquer transcrito. Este sistema gera a necessidade de distinguir o que é endógeno do que é exógeno. Para tal, crê-se na existência de um mecanismo de memória que protege os genes que devem ser expressos. Esse mecanismo é mediado pela única proteína Argonauta que é essencial em *C. elegans* – CSR-1 ("Chromosome Segregation and RNAi Deficient-1). CSR-1 também utiliza 22G, no entanto estes não são formados pelas proteínas "mutator", e crê-se que a sua actividade iniba o reconhecimento por PRG-1, impedindo assim o silenciamento de genes essenciais. No entanto, muito pouco se sabe a dinâmica entra as vias de silenciamento e a via

de activação (CSR-1).

Durante o trabalho desenvolvido nesta tese, identificámos dois factores envolvidos no silenciamento genético mediado por PRG-1: PID-1 e PID-2. PID-1 está envolvido na estabilização e talvez no transporte de percursores de piRNAs antes de estes serem processados e incorporados na PRG-1. Na ausência de PID-1 não existe produção de piRNAs. Por sua vez, PID-2 parece actuar após a via secundária mediada por proteínas “mutators”. Mutantes para PID-2 conseguem estabelecer um silenciamento secundário iniciado por PRG-1, mas falham em estabelecer silenciamento epigenético (RNAe), pelo que pensamos que seja o primeiro factor identificado que conecta os dois mecanismos.

Por fim, mostramos a via dos piRNAs é essencial para restabelecer o normal silenciamento das proteínas mutator. Quando reactivamos a via dos mutator na ausência de PRG-1, verificamos graves defeitos de proliferação das células germinais. Estes defeitos são acompanhadas por um desvio de pequenos RNA 22G da via de activação para a via de silenciamento, o que resulta no silenciamento de genes essenciais para o desenvolvimento de células germinais de *C. elegans.* Os nossos resultados abrem uma nova perspectiva sobre o equilíbrio entre silenciamento e activação de genes, e sugerem um modelo no qual a actividade de silenciamento é inibida não só por 22Gs herdados de ambos os progenitores, mas também por piRNAs maternais.

Summary

One of the challenges faced by germ cells is to repress molecular parasites, such as transposons and retro-viruses that when left unchecked can lead to fertility problems. Organisms have developed several mechanisms to accomplish this task, such as the use of Piwi-interacting RNAs (piRNAs) to identify and silence foreign sequences. In *C. elegans*, the non-essential PIWI Related Gene-1 (PRG-1) uses a repertoire of more than 30000 unique piRNAs to target most, if not all, RNA transcripts. Recognition by PRG-1 triggers a secondary RNAi response mediated by mutator genes, 22G endo-siRNAs, and the Argonaute proteins HRDE-1 and WAGO-1. By a poorly understood mechanism, mutator activity can in turn trigger a nuclear RNAi response that becomes independent of PRG-1. This state, known as RNAe, is heritable and requires HRDE-1 and the mutator genes. To avoid silencing its own genes, *C. elegans* is able to actively prevent silencing of specific genes. This gene-activating pathway is mediated by CSR-1, but the exact molecular mechanisms remain to be determined.

During the work conducted for this thesis, we performed a forward genetic screen for piRNA Induced Silencing Defective mutants, where we isolated two novel factors: PID-1 and PID-2.

PID-1 is involved in piRNA biogenesis and *pid-1* mutants show a depletion of mature piRNAs coupled to a relative accumulation of piRNA precursors, suggesting that *pid-1* in involved in transport or processing of piRNA precursors.

pid-2 mutants show normal levels of piRNA and also show secondary 22G resulting from PRG-1 targeting. Despite having an apparent silencing response triggered by PRG-1, *pid-2* mutants do not fully silence the sensor and are resistant to RNAe onset. Therefore we propose that *pid-2* may be the link between secondary mutator activity and RNAe onset.

Finally, while studying how the transgenerational memory of RNAe and the piRNA pathway interact, we found that maternally provided piRNAs are essential to properly re-establish mutator activity. Re-activation of mutator activity in the absence of PRG-1 results in severe defects in germline proliferation. These defects require HRDE-1 and are accompanied by a shift in the balance between gene-activating and gene-silencing pathways. In summary, our results demonstrate that maternal piRNAs in *C. elegans* are required to initiate transposon-targeting 22G-RNA pathways and to prevent ectopic activity of HRDE-1-mediated silencing. These results shed light on

the yet elusive balance between gene-silencing and gene-activating mechanisms and support a model in which HRDE-1 activity is prevented from targeting active genes through both maternally inherited piRNAs and parentally transmitted 22G RNAs.

Chapter 1

General Introduction

RNA, the change of a paradigm

Less than a decade after the discovery of the DNA structure (Watson and Crick, 1953) Francis Crick formally proposed two general principles that are hallmarks of molecular biology: The Sequence Hypothesis and the Central Dogma (Crick, 1958). The first states that “the specificity of a piece of nucleic acid is expressed only by the sequences of its bases, and that this sequence is a (simple) code for the amino acid sequence of a particular protein.” The second defends that the flow of information between nucleic acids and protein is unidirectional and that once information “has passed into protein it cannot get out again”. Although the Central Dogma does not exclude the flow of information from RNA to DNA, later confirmed by the discovery of an RNA dependent DNA polymerase (Baltimore, 1970; Temin and Mizutani, 1970), it did wave RNA aside as a mere intermediate in the flux of information. This paradigm slowly changed during the following decades with the notion that mRNA secondary structure could regulate its own transcription (Lee and Yanofsky, 1977) and that it possesses catalytic activity (Kruger et al., 1982; Stark et al., 1978). However, it was not until almost 40 years later that RNA gained importance as a key regulatory element. It was shown that chromosome imprinting is mediated by a long non coding RNA transcript (Lee et al., 1996) and most notably that double stranded RNA could mediate gene expression, through a process called RNAi (Fire et al., 1998). The discovery by Craig Mello and Andrew Fire that double stranded RNA can regulate gene expression re-shaped molecular biology research by opening the door for new tools and therapeutics.

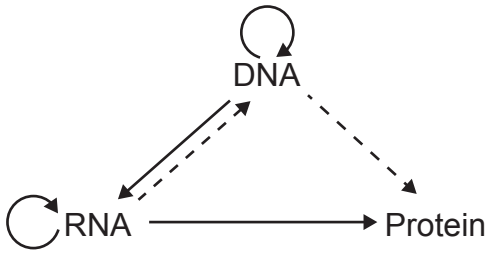


Figure 1. The Central Dogma. Representation of the central dogma of molecular biology, published by Francis Crick in 1970. It shows “the situation as it seemed in 1958”, solid arrows represent probable transfers of information and dotted arrows represent possible transfers of information. The absent arrows represent impossible flow of information (Crick, 1970).

Introducing RNA interference

RNA interference (RNAi) can be defined as a process by which an Argonaute protein recognizes a target transcript by means of a small RNA cofactor. Upon target recognition, Argonautes mediate the recruitment of other factors that initiate downstream events. Most of the times these events lead to down regulation of gene expression, either by degrading the targeted transcript or by inhibiting transcription, in a process generally known as gene silencing.

This phenomenon was first observed by Izant in 1984 (Izant and Weintraub, 1984) when he saw that transfecting cells with antisense RNA would reduce gene expression of a particular gene, at the time this was thought to be the result of annealing between the exogenous antisense RNA and the sense transcript. Later Andrew Fire and Craig Mello, while investigating the requirements for this RNA interference phenomena, observed that worms exposed to double stranded RNA, with the same sequence as the *unc-22* gene, show an *unc-22* mutant phenotype (Fire et al., 1998). Gene silencing triggered by double stranded RNA was not only much stronger than previously reported by Izant but also transmitted through generations after the initial RNA treatment. Simple annealing between endogenous sense transcript and exogenous anti-sense RNA could not explain this strong silencing effect; therefore, Fire and Mello suggested the existence of and underlying mechanism mediating RNA interference which would have a biological purpose, probably for physiological gene silencing. This mechanism has since been described in several eukaryotes such as *Drosophila* (Kennerdell and Carthew, 1998), tobacco plant (Waterhouse et al., 1998), zebrafish (Wargelius et al., 1999), mammalian cells (Elbashir et al., 2001) and fission yeast (Volpe et al., 2002).

In this introduction I will briefly describe the current knowledge of the mechanisms underlying three RNAi related pathways that are present in most organisms: the micro-RNA pathway, the siRNA pathway and the piRNA pathway.

The Argonaute

The central element in RNAi is the Argonaute protein, a highly conserved RNA binding protein with family members found across all eukaryotes, with the exception of *S. cerevisiae* (Drinnenberg et al., 2009). Argonaute proteins bind to small (20-26 nucleotides long) single stranded RNAs, or in some cases to single stranded

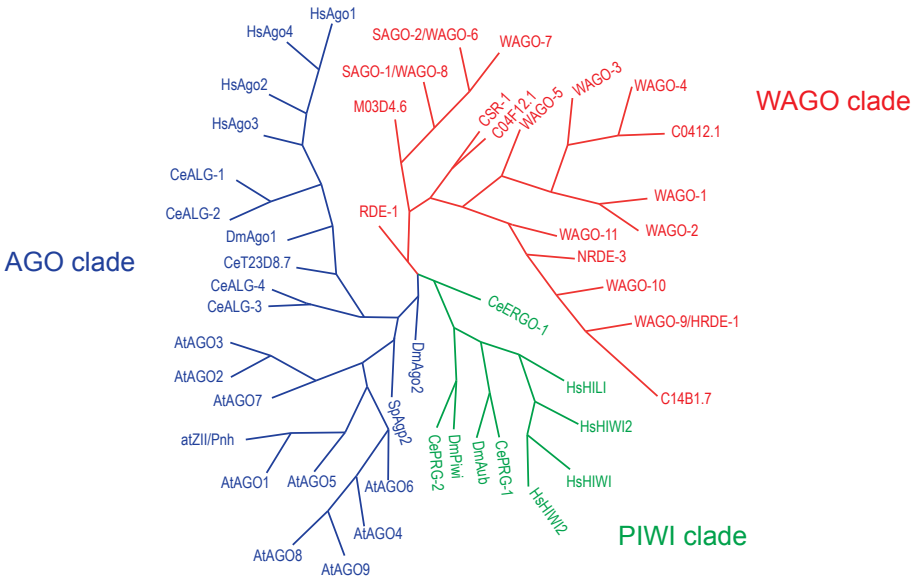


Figure 2. Argonaute clades. Phylogenetic tree of Argonaute proteins from *H. sapiens*, *D. melanogaster*, *C. elegans*, *S. pombe* and *A. thaliana*. In blue is represented the AGO clade, in green the PIWI clade and in red the worm specific WAGO clade. Adapted from Yigit et al., 2006.

precursors that are subsequently trimmed, and use these as guides to select their target transcripts through base-pair complementarity. Upon target recognition they promote a myriad of downstream processes which can vary depending on each specific Argonaute.

Argonautes contain three domains: PAZ (Piwi/Argonaute/Zwille), MID (middle) and PIWI (P-element induced wimpy testis). The PAZ domain forms a pocket that binds to the 3' end of the small RNA, the Mid domain provides a pocket for anchoring the 5' end of the small RNA and confers specificity between the utmost 5' nucleotide of a small RNA and certain Argonautes. The PIWI domain's tertiary structure resembles an RNase H enzyme and possesses a DDH catalytic motif, being able to slice the targeted transcript. However, contrary to what was initially thought, not all Argonautes require catalytic activity to function. In fact, several Argonaute proteins in C. elegans lack key catalytic residues at the DDH motif (Yigit et al., 2006) and in mouse AGO1, AGO3 and AGO4 are also catalytically dead (Liu et al., 2004). Other Argonautes such as Miwi2 in mouse, PRG-1 and RDE-1 in worms have an intact catalytic motif but do not require catalytic activity for their function (Bagijn et al., 2012; De Fazio et al., 2011; Steiner et al., 2009). The C-terminal of the PIWI domain

bends over the Mid domain further stabilizing the binding of the small RNA's 5' end (Elkayam et al., 2012).

Most organisms encode several different Argonautes, each specific for a certain tissue or function. Sequence conservation analysis of eukaryotic Argonautes reveals the existence of four clades of Argonautes (Swarts et al., 2014). The AGO clade includes Argonautes closely related to the Arabidopsis AGO-1. These Argonautes are generally ubiquitous and are involved in micro-RNA and siRNA pathways. The PIWI clade is composed by Argonautes close to the Drosophila PIWI, these are mostly germline specific and are involved in the silencing of transposons and other exogenous sequences. The Trypanosoma Ago clade encompasses Argonaute proteins found in protozoa (Swarts et al., 2014). Argonautes from this clade possess a divergent PAZ domain and are found to be involved in transposon silencing (Djikeng et al., 2001; Garcia Silva et al., 2010). The fourth and last group is the WAGO clade which is composed by Argonautes specific to nematodes. Argonautes within this clade are involved in several distinct pathways that sometimes have opposing roles. Some have been shown to mediate silencing of their targets while others have the opposite effect. Remarkably, most WAGOs also lack the conserved DDH motif in the PIWI domain, indicating that they have lost their catalytic activity. It's not clear why *C. elegans* encodes so many unique Argonautes, but it is an indication of how complex the interaction between these pathways can be, a topic that will be discussed later.

micro-RNAs

miRNAs are a class of highly conserved 20-24 nucleotide long small RNAs that have an important role in post transcriptional regulation of gene expression. In short, a double stranded RNA precursor is processed by Dicer, an RNase III, into 20-24 nucleotides long double stranded RNAs that in turn are loaded onto the microRNA-Induced Silencing Complex (mRISC), which contains an Argonaute protein. That Argonaute guides the RISC to its target, leading to mRNA destabilization or translational inhibition. The micro-RNA pathway modulates mRNA activity and is an important process that fine-tunes mRNA stability and translation. Micro-RNAs are encoded throughout the genome and transcribed by RNA Polymerase II (Pol II). Upon transcription, the transcript folds itself into a double-stranded molecule due to intramolecular base-pairing, creating an hairpin loop. This RNA precursor can be transcribed as an individual gene (pri-miRNA) or be encoded in an intron (mirtron).

pri-miRNAs are processed by the nuclear microprocessor where Drosha releases the hairpin (pre-miRNA). Mirtrons are Drosha independent and processed during splicing. Pre-miRNAs are then transported to the cytoplasm by Exportin 5 and loaded into Dicer which cleaves the hairpin's loop, generating a double stranded RNA composed of the mature miRNA and its complementary strand, usually referred as miRNA*. The mature miRNA is then loaded into a complex named miRISC where it directly binds to an Argonaute from the AGO clade. The miRNA* is usually degraded although it might be functional in some cases. In Drosophila, micro-RNAs can also be processed independently of Dicer and loaded directly to AGO-2.

miRISC identify their target through base pair complementary between the ago bound sRNA. In animals, full complementarity between miRNA and the target transcript is not required and there are usually one or more mismatches between them. In most cases, only bases from the position 2 through position 8, a region also known as seed sequence, require full complementarity with the target transcript. The recognition sites of miRNA are usually at the 3' UTR of transcripts and the mechanism of silencing can either be translational repression or mRNA decay. During translational repression, the miRSC inhibits translation initiation impairing eIF4E-CAP recognition and inhibiting ribosomal complex formation, whereas during mRNA decay, miRISC interacts with the CCR4-NOT and PAN2-PAN3 facilitating dead enylation of the poly(A) tail, resulting in transcript destabilization.

siRNAs

Mechanistically, the main difference between the micro-RNA pathway and the siRNA pathway is the source of the double stranded RNA precursor. In endogenous siRNAs (endo-siRNAs), precursors are long double stranded RNA molecules that result from inter-molecular interaction between different transcripts, either from bidirectional transcription at the same locus or between transcripts from different loci. In mouse oocytes and mouse embryonic stem cells, endo-siRNAs precursors can also be long hairpins that originate from a genomic locus with tandem, inverted short interspersed nuclear elements (SINEs). Endo-siRNA double stranded precursors can also be formed by the activity of RNA directed RNA polymerases (RdRP) present in plants, *S. pombe* and *C. elegans.* However, the most commonly described form of siRNA is not endogenous, but exogenously provided through a purely artificial system and is a tool extensively used to induce gene silencing

in different organisms. endo-siRNA do not undergo nuclear processing, instead they are directly processed by Dicer, that slices the longer double stranded RNA precursor. In case of metazoan, exogenous siRNAs are often introduced as shRNAs that structural similar to pre-miRNA and therefore bypass the Microprocessor and are processed directly by DICER. The guide strand, the equivalent of a mature miRNA, is then loaded to the siRISC whereas the passenger strand, the equivalent of the miRNA*, is destabilized and degraded.

piRNAs

PIWI-interacting RNAs (piRNAs) are a class of small RNAs found in animals and that are associated with the PIWI clade of Argonaute proteins (Cox et al., 1998; Lin and Spradling, 1997). PIWI Argonautes are predominantly expressed in germ cells and supporting somatic tissues where their primary role is to protect the genome against transposons and other non-self molecular species (Malone and Hannon, 2009). piRNA function is essential and impairment of the PIWI pathway results in severe defects in germline development and sterility (Carmell et al., 2007; Das et al., 2008; Houwing et al., 2007; Lin and Spradling, 1997). Biogenesis of piRNAs is different from microRNAs and siRNAs, they result from a longer single strand precursor and are therefore, not dependent on Dicer activity (Brennecke et al., 2007; Houwing et al., 2007; Vagin et al., 2006). In most organisms, piRNA precursors are transcribed from specific loci composed of remnants of transposons and other repetitive sequences, named piRNA clusters. Precursors are transcribed by Pol II as long single stranded transcripts, up to more than 100K base-pairs, and then processed into several piRNAs (Aravin et al., 2006, 2007; Brennecke et al., 2007; Girard et al., 2006; Houwing et al., 2007; Lau et al., 2006).

In fly uni-strand piRNA clusters require the transcription factor Cubitus interroutus (Goriaux et al., 2014) and dual-strand clusters depend on the Rhino-Deadlock-Cutoff complex (Klattenhoff et al., 2009; Mohn et al., 2014; Pane et al., 2011; Zhang et al., 2014). After transcription, fly piRNA precursors are exported by UAP56 and Vasa to the nuage, a cytoplasmic electron dense structure, where they undergo primary processing. This primary pathway involves the cleavage of the long precursor into smaller single stranded RNAs by Zucchini (Ipsaro et al., 2012; Nishimasu et al., 2012) and loading into Piwi/Aubergine (Aub) with the help of Shutdown (shu) and Hsp83 (Izumi et al., 2013; Olivieri et al., 2012; Preall et al.,

2012). Piwi/Aubergine binding is biased towards fragments with an uracyl at the 5' prime position. After loading into Piwi, the 3' prime overhang is trimmed by a yet unknown 3'-5' exonuclease (Trimmer) to the length of the mature piRNA (Kawaoka et al., 2011).The last step consists in the 2'-O-methylation of the last base (3' end) by the methyl-transferase Hen-1 (Horwich et al., 2007; Saito et al., 2007), which protects the mature piRNA from urydilation that would eventually destabilize and degrade it.

Upon synthesis through the primary piRNA pathway, fly piRNAs are carried by Aub to a secondary piRNA pathway, also called ping-pong amplification, in which transcripts homologous to the piRNA sequence are degraded in a process that also produces more piRNAs, creating a positive feedback loop. Aub, loaded with a primary piRNA, recognizes a transcript by base-pair homology and cleaves it. Ago3 (another Argonaute from the PIWI clade) is then recruited to the 3' prime fragment of the cleaved transcript. Ago3 is now loaded with a secondary piRNA precursor and undergoes the same process of trimming and methylation as Piwi/Aub in the primary pathway. Ago3 is then able to identify a new target, slice it and recruit Aub in a similar fashion, closing the cycle and creating a positive feedback loop that continues to degrade transcripts homologous to the ones that created the primary piRNAs.

This strategy, adopted by *Drosophila*, allows the silencing in trans of transcripts whose sequence is represented in the piRNA clusters. However, new invading molecular parasites are invisible to this protection mechanism since they are not represented in the piRNA clusters. One potential solution resides on the very replicative nature of transposable elements, hoping that at some point one insertion lands in these "traps" that are constantly producing primary piRNAs. Once one of those events occurs, the secondary pathway ensures the silencing of the remaining foreign copies. An example of this mechanism is the adaptation to the P element transposon in fly. In drosophila, piRNAs are maternally provided and paternally inherited P element transposons escape piRNA mediated silencing. Hybrids where the P element is paternally provided are sterile due to germline DNA damage and transposon mobilization. However their sterility is recued once the P element in integrated into a piRNA cluster and piRNAs targeting the P element are made (Khurana et al., 2011).

piRNAs in *C. elegans*

Like in other organisms, identification of invading species is carried by a Piwi like Argonaute, PRG-1 (Piwi Related Gene 1) (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008), through the use of piRNAs, but both the biogenesis of these piRNAs and the mechanism by which they lead to silencing are quite different. The structure of a worm piRNA (21U small RNA) is similar to other organisms, they have a strong consensus to a U at the 5' end and their last nucleotide is 2-O-methylated (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012), they are however shorter and with a fixed length – 21 nucleotide long, contrary to the fly and mouse piRNAs that range from 26 to 30 nucleotide long (Aravin et al., 2006, 2008; Girard et al., 2006). *C. elegans* encodes a second Piwi like Argonaute (PRG-2) that seems to have little or no function in the piRNA pathway (Batista et al., 2008; Das et al., 2008). Despite recent efforts to describe it, worm piRNA biogenesis is still largely unknown. Contrary to other organisms, *C. elegans* piRNA are individually transcribed from discrete genetic units in the form of 28-35 nucleotide long precursors that are further processed into the mature 21U piRNA. There are more than 30000 identified piRNAs, grouped into 2 classes that differ in how piRNA precursors are transcribed (Gu et al., 2012). Type I piRNAs, the most abundant ones, share the 8 nucleotide long consensus motif "CTGTTTCA" (Ruby motif) upstream each locus (Ruby et al., 2006) which is recognized by several members of the Forkhead family of transcription factors (Cecere et al., 2012) and are essential for precursor transcription. Type I piRNA loci are enriched in two clusters on chromosome IV. PRDE-1 (piRNA Defective 1) and SNP-4 (a DNA binding protein) co-localize to these clusters and are involved in generation the piRNA precursors. (Kasper et al., 2014; Weick et al., 2014). Type II piRNAs loci do not have an apparent upstream motif and are spread throughout the genome, being probably a bi-product of early Pol II termination (Gu et al., 2012). In both types, precursors are products of Pol II transcription and contain a 5' 7-methylguanylate cap. Interestingly, transcription starts precisely 2 nucleotides upstream of the 5' U of the mature piRNA at the YRNT motif and extends up to 15 bases after the 3' nucleotide of the mature piRNA (de Albuquerque et al., 2014; Gu et al., 2012). A recent RNAi screen for genes involved in piRNA biogenesis identified several factors named TOFUs (Twenty-one fouled up). The precise role of these factor is still unknown, however analysis of the small RNA populations revealed that depletion of TOFU-3, TOFU-4 and TOFU-5 levels

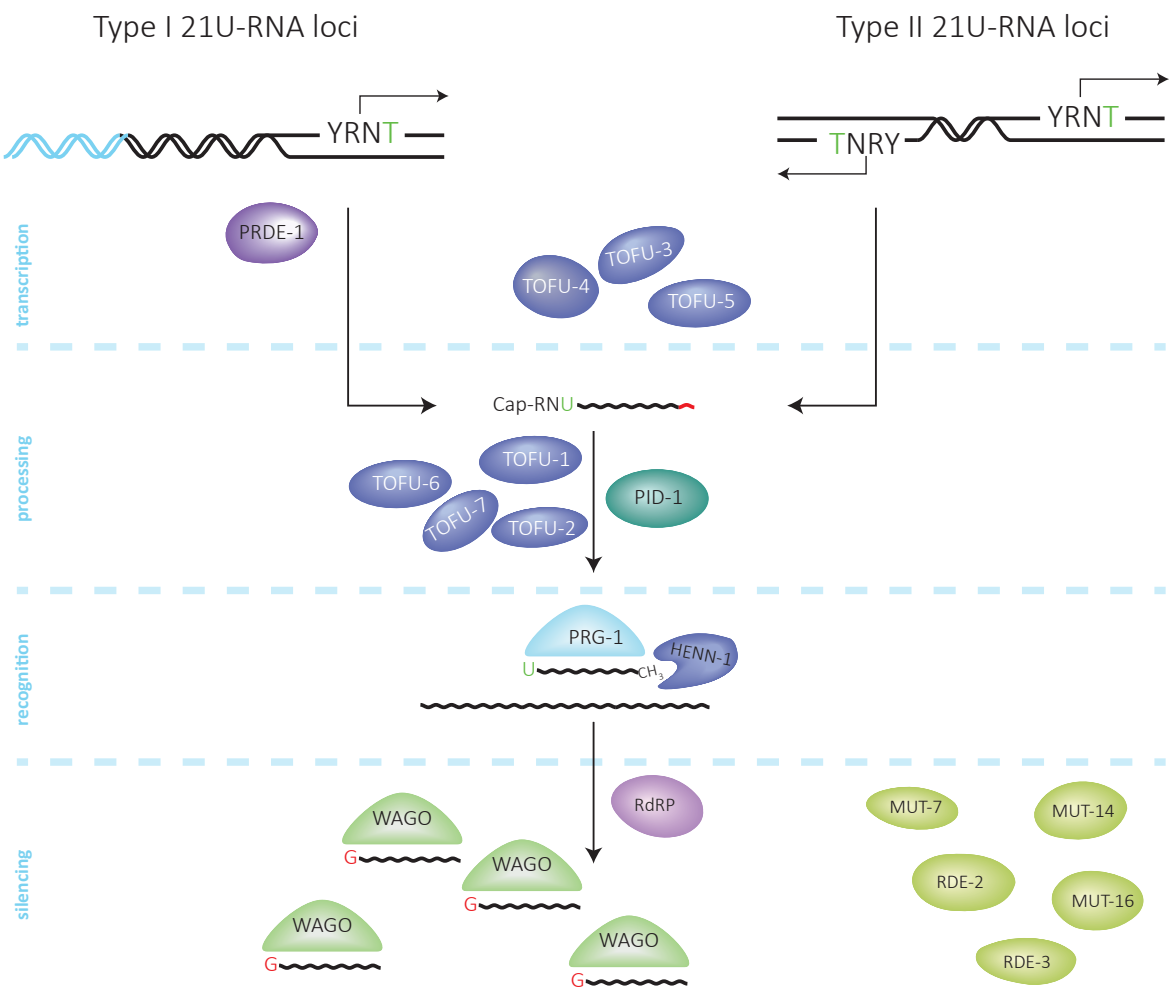


Figure 3: piRNA pathway in *C. elegans*. In *C. elegans*, piRNAs precursors are transcribed as independent short transcripts of 26~30 nucleotides either from DNA clusters on chromosome II and IV (Type-I 21U-RNAs) or , in case of the less abundance Type-II 21U-RNAs, derive from promoters of Pol II genes spread throughout the genome. After transcription, precursors are processed by a still unknown mechanism into a 21 nucleotides long small RNA with a bias for uracil at the 5' end (21U RNA) and loaded into the Piwi argonaute PRG-1. Further modifications include methylation of the 3' end by the RNA methyltransferase HENN-1. Recent studies identified several factors involved in precursor transcription/processing, but their precise role is still unknown. PRG-1, bound to its piRNA (21U) cofactors, is then involved in the identification of non-self-sequences and triggers a secondary 22G mediated response that degrades the targeted transcript and may initiate RNA-induced epigenetic silencing (RNAe).

results in a significant reduction of piRNA precursors (Goh et al., 2014). TOFU-3 is a SUMO-related protease whose mammalian homolog SENP7 regulates HP1 localization to pericentric heterochromatin (Maison et al., 2012) whereas TOFU-5 has a SANT-like domain that can participate in chromatin remodeling (Boyer et al., 2004), suggesting that these two factors are involved in precursor transcription, yet it is unknown if they affect one or both types of precursors.

After transcription, piRNA precursors are likely to be exported from the nucleus

to perinuclear granules called p-granules (Strome and Wood, 1982) (an electron dense structure equivalent to the nuage in other organisms) where PRG-1 is localized. In the p-granules, precursors are processed into to mature piRNAs by removing the 2 5' bases and 3' trimming, finalized by 2-O-methylation by the HEN1 homologue HENN-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). In a forward genetic screen we identified PID-1 (pi-RNA Induced silencing Defective) (described in detail in chapter 2) as a novel factor involved in the stability and/or transport of both types of precursors (de Albuquerque et al., 2014). *pid-1* mutants show an overall decrease of both precursors and mature piRNAs, but a relative increase in precursors in relation to mature piRNAs. Since the overall molecular structure of both precursors and mature piRNAs is still maintained, we suggest that PID-1 is somehow involved either in transport or stability of precursors. We also found that PID-1 is associated with TOFU-6/MEL-47 which contains a RNA Recognition Motif (RRM) (unpublished). Knockdown of *tofu-6* and *tofu-7* results in a decrease in mature piRNAs while keeping normal levels of precursors, suggesting that they act at the same levels as PID-1 (Goh et al., 2014). Mutants for other two factors, TOFU-1 and TOFU-2, show a clear accumulation of piRNA precursors and a depletion of mature piRNAs comparable to *prg-1* mutants. Both contain an ATP-binding site suggesting that they are involved in processing the precursors in a ATP-dependent process (Goh et al., 2014). 2-O-methylation of the piRNA is thought to occur after loading to PRG-1 and only mildly affects piRNA function (Kamminga et al., 2012). Upon formation, the PRG-1/21U complex finds its targets by incomplete base-pair complementarity (Bagijn et al., 2012). Unlike other organisms, C. elegans PRG-1 does not slice its targets (Bagijn et al., 2012). Instead, it recruits an RNA-directed RNA polymerase (RdRP) that makes secondary siRNAs (22G siRNAs) from the targeted transcript. This secondary amplification pathway is mediated by worm specific Argonautes (WAGO-1 and HRDE-1) and mutator genes (Bagijn et al., 2012). It is not known how the targeted transcript is silenced, but after continued targeting by PRG-1, a transcript becomes transcriptionally silenced in a self-maintained mechanism called RNAe (RNA-induced epigenetic silencing) (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Once RNAe has been established there is an increase of H3K9 tri-methylation (H3K9me3) at the targeted locus, and silencing becomes independent of PRG-1but still mediated by the nuclear Argonaute HRDE-1 and mutator genes (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012). At the small RNA level, initial targeting by PRG-1

results in the build-up of a 22G population antisense to the transcript and near the piRNA recognition site (Bagijn et al., 2012; Lee et al., 2012), this population shifts away from the piRNA recognition site towards the 5' end of the transcript once RNAe is established (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). The establishment of this PIWI independent silencing may explain why in the absence of PRG-1 all but one transposon family (*Tc3*) remain silenced (Das et al., 2008).

The strategy by which C. elegans identifies new molecular parasites is also different from other studied models. Whereas in other organisms all transcripts are licensed for translation, unless represented in the piRNA clusters, *C. elegans* adopted the opposite strategy to silence foreign sequences. It labels all transcripts as "bad" and uses a licensing list to ensure that the "good" ones are protected from silencing. Due to the existence of more than 30000 piRNA that recognize their targets by partial pairing, virtually any transcript can be targeted for silencing (Bagijn et al., 2012; Lee et al., 2012). On the one hand this is a powerful strategy to silence new foreign sequences, on the other it also leads to silencing of the worm's endogenous genes. To face this it has been proposed that C. elegans transfers the information of which genes are being expressed from one generation to the other, by means of another RNAi pathway mediated by the essential Argonaute CSR-1 (Conine et al., 2013; Seth et al., 2013; Wedeles et al., 2013).

WAGOS, RdRPs, Mutators and the secondary pathway

Besides the AGO and PIWI classes of Argonautes, the *C. elegans* genome encodes an extra 18 unique Argonautes. These Argonautes form the WAGO clade (Worm Argonautes) and are involved in the secondary siRNA pathway. In the worm, both piRNA and siRNA pathways are inefficient at silencing and require a secondary amplification/silencing step. These secondary siRNA pathways are mediated by the WAGOs and a particular class of small RNAs called 22G siRNAs. 22G RNAs, which as the name indicates, are 22 nucleotide long and have a strong bias to a guanidine at the 5' end. The 22G siRNAs are product of RdRP activity and are tri-phosphorylated at their 5' end, which suggests a de-novo synthesis of each 22G siRNA instead of slicing a longer precursor. Although using structurally identical siRNAs, different WAGOS may have different functionalities with some having opposing roles.

C. elegans has yet another particularity with respect to other animals: it requires RdRP activity for functional RNAi. There are four RdRP encoded by the

worm genome: *rrf-1*, *rrf-2*, *rrf-3* and *ego-1*. RRF-1 is required for the secondary siRNA pathway, is ubiquitously expressed and can be replaced by EGO-1 in the germline. EGO-1 is the only essential RdRP and is mainly involved in the CSR-1 pathway. RRF-3 is required for a primary siRNA pathway mediated by ERGO-1 and ALG-2/3 and there is no known function of RRF-2.

A key component of the secondary siRNA pathway is the mutator complex. Mutator (mut) genes were the first identified as required for transposon silencing in the germline (Collins et al.). Mutations on these genes show a high increase of spontaneous mutations as a result of transposon activity in the germline (Collins et al.). All six characterized mutator genes (*mut-2/rde-3*, *mut-7*, *mut-8/rde-2*, *mut-14/smut-1*, *mut-15* and *mut-16*) are temperature-sensitive sterile and show a high incidence of males (*him*) phenotype, probably due to chromosomal nondisjunctions (Ketting et al., 1999; Vastenhouw et al., 2003). With the exception of *mut-14*, they are ubiquitously expressed and required for both somatic and germline RNAi. In the germline they aggregate in structures that surround the p-granules and are called mutator foci. These structures also co-localize with RdRP RRF-1 suggesting that secondary 22G siRNA biogenesis might be done in these “compartments” (Phillips et al., 2012). The 3’-5’ exonuclease *mut-7*, a homolog of the Werner Syndrome Helicase and RNAse D (Ketting et al., 1999), is present both in the nucleus and in the cytoplasm, where it associates with the mutator foci (Phillips et al., 2012) and its thought to be directly involved in the degradation of the target mRNA (Ketting et al., 1999). The cytoplasmic fraction interacts with *mut-8/rde-2*, (Tops et al., 2005) and in the absence of *mut-8/rde-2*, *mut-7* is no longer enriched at the mutator foci (Phillips et al., 2012). *mut-14* and its paralog *smut-1* (Synthetic Mutator-1) are RNA helicases that have overlapping functions (Phillips et al., 2014; Tijsterman et al., 2002). Contrary to other mutator genes, *mut-14* and *smut-1* are only required for germline RNAi (Phillips et al., 2014). The close homology of MUT-14 with the D. melanogaster’s Vasa, which transports piRNA precursors in the nuage, suggest that *mut-14/smut-1* may be involved in the transport of mRNAs from the p-granules into the mutator foci. If true, their role would not be necessary in the soma since p-granules are only present in the germline. MUT-16 is only conserved among nematodes and contains glutamine/asparagine (Q/N)-rich domains which are normally associated with both protein self-aggregation and protein/protein interactions (Michelitsch and Weissman, 2000). mut-16 mutants are the only mutator mutants that fail to form the mutator foci surrounding the p-granules, suggesting a scaffold function (Phillips

et al., 2012; Vastenhouw et al., 2003; Zhang et al., 2011). RDE-8, a NYN domain ribonuclease, has recently been show to transiently bind to and cleave mRNA targeted by exo-siRNA and to be required for the accumulation of 22G siRNAs. RDE-8 localizes to mutator foci and directly interacts with *mut-15*, *mut-16* and *mut-2/rde-3*, a nucleotidyltransferase whose polymerase activity is essential for RNAi (Chen et al., 2005; Tsai et al., 2015). Upon mRNA cleavage by RDE-8, *mut-2/rde-3* urydilates the targeted mRNA which in turn leads to *rrf-1* recruitment and 22 siRNA production (Tsai et al., 2015). The function of *mut-8/rde-2* and *mut-15* is still unknown.

Exogenous siRNA

Whereas in other organisms, exo-siRNA mediated silencing uses the micro-RNA pathway mechanism, in *C. elegans* exogenous double stranded RNA (dsRNA) is processed by a primary siRNA pathway mediated by RDE-1 (Tabara et al., 1999), an Argonaute from the WAGO clade. Like in other organisms, DICER (DCR-1) binds to long dsRNA in the cytoplasm and dices it into smaller short dsRNAs. However, instead of being loaded into the AGO-1 homologs ALG-1 and ALG-2 (Grishok et al., 2001), these DCR-1 products are loaded into RDE-1 (Tabara et al., 1999). DCR-1 is in a complex with RDE-1, DRH-1 (Dicer Related Helicase 1) and RDE-4 (Tabara et al., 2002), a RNA binding protein that has high affinity for long double stranded RNAs (Parker et al., 2006), and it is believed that recognition of the dsRNA by RDE-4 triggers the recruitment of DCR-1 and the remaining components. Loading of the dsRNA into RDE-1 instead of the ALG-1/2 can also be due to a higher affinity of RDE-1 towards perfectly or near perfect complementary dsRNAs (Steiner et al., 2007). RDE-1 slicing activity is required to remove the passenger strand from the dsRNA but not for target degradation (Steiner et al., 2009). Upon passenger strand removal, RDE-1 recognizes its target mRNA and transfers it to the to the mutator foci in a process mediated in part by the DEAD box helicase RDE-12 (Shirayama et al., 2014; Yang et al., 2014) ,RDE-10 and the zinc finger RDE-11 (Yang et al., 2012; Zhang et al., 2012). At the mutator foci, RDE-8 is recruited to the target, initiating targeted degradation and secondary 22G siRNA response as described previously.

Endogenous siRNAs

Besides the exo-siRNAs and secondary 22G endo-siRNAs, the nematode

C. elegans also has primary endo-siRNAs which are structurally different from the secondary siRNAs (Ruby et al., 2006). They are 26 nucleotide long and have a strong bias towards a guanidine at the 5' position (hence the name 26G). 26G siRNAs are the result of RRF-3 activity that, along with other co-factors, generates a double stranded RNA precursor from the targeted transcript. This double stranded precursor is then diced by DCR-1 and loaded into one of two Argonautes, ERGO-1 (PIWI clade) or ALG-3/ALG-4 (AGO clade) which in turn have distinct functionalities. Several of the factors involved in 26G biogenesis were first identified as ERI genes (Enhanced RNAi). Initially it was proposed that this effect was due to suppressed degradation of siRNAs, since ERI-1 contains both a nucleic acid binding SAP and a 3'-5' exonuclease domain and thought to be responsible for siRNA degradation (Kennedy et al., 2004). However, later studies shown that ERI-1 is actually involved in the biogenesis of 26G along with ERI-3, ERI-5 (a Tudor protein), RRF-3, RDE-4 and DCR-1 (Duchaine et al., 2006; Thivierge et al., 2011; Vasale et al., 2010). The increased sensitivity to exo-siRNAs is due to competition between endo and exo siRNA pathways for shared components during both 26G biogenesis and downstream 22G siRNA amplification (Vasale et al., 2010).

Based on their association with a specific Argonaute and their expression pattern, 26G endo-siRNAs can be separated into two groups. Ergo-1 is expressed in oocytes and embryos and mutants show a depletion of both 26G and secondary 22G siRNAs against a subset of genes. Many of those targeted genes appear to be ancient gene duplications, suggesting the function of this pathway may be to buffer expression of rapidly expanding gene families (Vasale et al., 2010). Like in the exo-siRNA pathway there is a clear two-step mechanism for silencing ERGO-1 targeted genes: a primary pathway mediated by ERGO-1 with its associated 26G co-factors and the shared secondary response mediated by the mutator pathway, *rrf-1* and rde-8. Absence of *rrf-1*, *rde-8* and mutator genes leads to a depletion of 22G against ERGO-1 targets, but show normal levels of 26G endo-siRNAs (Phillips et al., 2014; Tsai et al., 2015; Vasale et al., 2010) whereas mutants for *ergo-1* and factors involved in 26G biogenesis, show a depletion of both 26G and corresponding 22G endo-siRNAs (Vasale et al., 2010). One unique structural feature of ERGO-1 bound 26G is their 3' 2-O-Methylation by HENN-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012), which seems to be essential for ERGO-1 activity. henn-1 mutants show increased urydilation levels at the 3' of ERGO-1 26Gs (Kamminga et al., 2012) , depletion of corresponding 22G endo-siRNAs and an

increased expression of targeted transcripts (Montgomery et al., 2012).

ALG-3 and ALG-4 are two paralogs from the AGO clade that are expressed in the germline during spermatogenesis and are essential for sperm development (Yigit et al., 2006). Although single mutants have a minor reduction in brood size, worms deficient in both these Argonaute proteins exhibit a drastic reduced brood size at elevated temperatures (Conine et al., 2010). Although both males and hermaphrodites double mutants produced a wild-type number of spermatids, these fail to undergo spermiogenesis (Conine et al., 2010). ALG-3/4 are associated with a subclass of 26G endo-siRNA, but in contrast to ERGO-1 bound siRNAs, these 26Gs are un-methylated and target more than 60% of sperm specific genes (Conine et al., 2013; Reinke et al., 2004). Like ERGO-1, ALG-3/4 also triggers 22G endo-siRNA production from targeted transcripts; however, these 22Gs seem to be independent of mutator activity. Since most genes targeted for by ALG-3/4 are essential for spermiogenesis, it seems unlikely that ALG-3/4 mediates gene silencing. In agreement with this, it has been shown that *alg-3/4* mutant males show decreased RNA levels, and the corresponding decrease in protein levels, of genes targeted by ALG-3/4 (Conine et al., 2013). It is proposed that ALG-3/4 siRNA pathway acts along with the CSR-1 pathway (discussed in the next topic) to promote expression of its targets.

The CSR-1 pathway and transcript licensing

The Argonaute CSR-1(Chromosome Segregation and RNAi Deficient-1) is unique among other Argonautes both for its function as for the fact that it is the only essential Argonaute in worms. It's mainly expressed in the germline where it is involved in chromosome segregation along with DHR-3 (an helicase) , EKL-1 (a tudor protein), EGO-1 and CDE-1 (a uracil-transferase) (Claycomb et al., 2009; van Wolfswinkel et al., 2009). Loss of any of these elements results in an improper alignment of the metaphase plate and kinetochores do not orient to opposing spindle poles during both meiosis and mitosis (Claycomb et al., 2009; van Wolfswinkel et al., 2009). CSR-1 mutant hermaphrodites are sterile but males are partially fertile at 20 ºC, probably due to a partial rescue by ALG-3/4 during spermatogenesis since csr-1;alg-3/4 triple mutants males are fully sterile (Conine et al., 2010). Like other WAGOs, CSR-1 interacts with 22G secondary siRNAs, however these siRNAs are not dependent on the activity of mutator genes (Phillips et al., 2014) and map anti-sense

to more than 4000 germline expressed genes (Claycomb et al., 2009). Moreover *csr-1* mutants show a slight decrease in expression of its targets, suggesting that it acts in the opposite way of other Argonautes by promoting transcript stability (Claycomb et al., 2009). This subclass of 22G siRNAs (CSR-22Gs) is structurally indistinguishable from the mutator dependent 22G that are found in the context of other WAGOS (WAGO-22G) and its still unknown what triggers their production. EGO-1 (Enhancer of glp-1) is the only RdRP responsible for their synthesis, contrary to the WAGOs-22G where EGO-1 can replace RRF-1. Another unique characteristic of this 22G population is their extensive 3' uridilation by CDE-1(van Wolfswinkel et al., 2009). Up to 40% of 22G that co-IP with CSR-1 are at least mono-uridilated, and loss of cde-1 phenocopies loss of csr-1. Interestingly, loss of CDE-1 also increases the CSR-1 22G levels, which suggests a very dynamic mechanism on which 22G bound to csr-1 would have a relative short half-life as a result of CDE-1 activity.

As mentioned before, in *C. elegans*, the piRNA pathway is capable of targeting every expressed transcript, which leads to the necessity of a counter balancing mechanism to avoid silencing of essential genes. Since CSR-1 targets germline expressed genes, without inducing silencing, this Argonaute protein is thought to be the key element in a mechanism by which information of previously licensed transcripts is carried through generations. In fact, it has been recently shown that coupling CSR-1 to a transgene prevents its silencing by piRNAs (Wedeles et al., 2013) and that CSR-1 is necessary when rescuing in trans an epigenetically silenced (RNAe) GFP transgene with a second non-RNAe GFP transgene (Seth et al., 2013) in a processed termed RNA activation (RNAa). How CSR-1 is able achieve this is still a mystery. One possible explanation is that CSR-1 is guided to "good transcripts" by CSR-1-22Gs where it deflects PRG-1/WAGO. One other explanation can be that the CSR-1 pathway degrades WAGO-22Gs that result from PRG-1 recognition, hence protecting the 'good' transcripts from silencing. It is still unclear how transcripts are selected to be targeted by the CSR-1 pathway, but it was proposed that CSR-1 acts downstream of ALG-3/4 in the male germline (Conine et al., 2013) to provide an epigenetic memory of what genes were being expressed in the previous generation.

Nuclear RNAi and epigenetic silencing

In addition to the effects at the mRNA level, siRNA pathways can also elicit changes in the epigenetic status of their targeted genes. This effect, called

Transcritptional Gene Silencing (TGS), was first described in plants where it was found that transgenic and viral RNA can trigger DNA methylation around the targeted sequence (Matzke et al., 1989; Wassenegger et al., 1994). TGS is was also reported in other organisms such was *S. pombe*, where Ago1 uses small RNAs derived from DNA repeats to promote gene silencing at pericentromeric region (Volpe et al., 2002), and *Tetrahymena*, where siRNAs were shown to induce localized chromatin modification during macronuclear development (Aronica et al., 2008; Malone et al., 2005; Nowacki et al., 2008).

In *C. elegans*, somatic exo-siRNA has been shown to induce deposition of the repressive mark H3K9me3 at the target loci (Burkhart et al., 2011) and to inhibit Pol II elongation (Guang et al., 2010). A forward genetic screen for nuclear RNAi components identified four factors that are involved in *C. elegans* TGS: NRDE-1, NRDE-2, NRDE-3 and NRDE-4 (Guang et al., 2008). NRDE-3 (Nuclear RNAi Defective -3) is the only factor that is strictly somatic and encodes an Argonaute from the WAGO clade that localizes to the nucleus (Guang et al., 2008). NRDE-3 is associated with 22G siRNAs and in mutants for either the ERGO pathway or mutator pathway, NRDE-3 fails to localize to the nucleus and remains in the cytoplasm. The change in localization due to absence of ERGO-1 derived secondary 22G siRNAs suggests that NRDE-3 is loaded with 22Gs at the mutator foci and then enters the nucleus where it mediates target silencing (Guang et al., 2008). Interestingly, in animals without endo-siRNA activity (*eri-1* or *ergo-1* mutants) NRDE-3 re-localizes to the nucleus upon treatment with exo-RNAi and in a target dependent manner, further reinforcing the idea that NRDE-3 transports secondary 22G to the nucleus (Guang et al., 2008). NRDE-2 is an evolutionary conserved protein that associates with NRDE-3 in the nucleus and is recruited to nascent transcripts by NRDE-3/siRNA complexes where it inhibits Pol II elongation (Guang et al., 2010). At the pre-mRNA, NRDE-2 recruits NRDE-1, a nematode specific factor that interacts with chromatin at the targeted locus. Chromatin interaction is dependent of yet another nematode specific factor NRDE-4, in *nrde-4* mutants NRDE-1,2 and 3 still localize to the pre-mRNA but NRDE-1 no longer associates with chromatin (Burkhart et al., 2011).

In the germline, exo-siRNA also leads to chromatin silencing at the targeted loci through the NRDE pathway. However since NRDE-3 is somatic, its role is replaced by the argonaute HRDE-1 (Heritable RNAi Deficient-1). HRDE-1 is a germline specific argonaute from the WAGO clade that localizes to the nucleus and it interacts with

22G siRNA that result from mutator activity (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). Contrary to NRDE-3 that mainly interacts with 22G from derived from ERGO-1 activity, HRDE-1 is bound to 22G mapping to non-ERGO-1 targeted genes, pseudogenes and transposons (Buckley et al., 2012; Shirayama et al., 2012). Moreover, germline TGS resulting from exo-siRNA is carried over generations, similar to what is observed with piRNA induced RNAe. In fact, both phenomena appear to share the same factors downstream mutator activity, since mutants for *hrde-1*, *nrde-1*, *nrde-2* and *nrde-4* loose exo-siRNA induced TGS and reset RNAe (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Therefore, the heterochromatin protein-1 (HP-1) homologs *hpl-1/2* and the histone methyl-transferases SET-25 and SET-32, described to be involved in maintenance of the RNA status of a GFP::H2B transgene (Ashe et al., 2012; Shirayama et al., 2012), may also be components of the NRDE pathway.

REFERENCES

De Albuquerque, B.F.M., Luteijn, M.J., Cordeiro Rodrigues, R.J., van Bergeijk, P., Waaijers, S., Kaaij, L.J.T., Klein, H., Boxem, M., and Ketting, R.F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28, 683–688.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442, 203–207.

Aravin, A. a, Hannon, G.J., and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318, 761–764.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. Mol. Cell 31, 785–799.

Aronica, L., Bednenko, J., Noto, T., DeSouza, L. V., Siu, K.W.M., Loidl, J., Pearlman, R.E., Gorovsky, M.A., and Mochizuki, K. (2008). Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in Tetrahymena. Genes Dev. 22, 2228–2241.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150, 88–99.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M.E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E. a (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science (80-.). 337, 574–578.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature 226, 1209–1211.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31, 67–78.

Billi, A.C., Alessi, A.F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., and Kim, J.K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. PLoS Genet. 8, e1002617.

Boyer, L.A., Latek, R.R., and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? Nat. Rev. Mol. Cell Biol. 5, 158–163.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089–1103.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489, 447–451.

Burkhart, K.B., Guang, S., Buckley, B. a., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mrna-associating factor links endogenous sirnas to chromatin regulation. PLoS Genet. 7.

Carmell, M. a, Girard, A., van de Kant, H.J.G., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev. Cell 12, 503–514.

Cecere, G., Zheng, G.X.Y., Mansisidor, A.R., Klymko, K.E., and Grishok, A. (2012). Promoters recognized by forkhead proteins exist for individual 21U-RNAs. Mol. Cell 47, 734–745.

Chen, C.C.G., Simard, M.J., Tabara, H., Brownell, D.R., McCollough, J.A., and Mello, C.C. (2005). A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr. Biol. 15, 378–383.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van Wolfswinkel, J.C., Chaves,

D. a, Shirayama, M., Mitani, S., Ketting, R.F., et al. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139, 123–134.

Collins, J., Saari, B., and Anderson, P. Activation of a transposable element in the germ line but not the soma of Caenorhabditis elegans. Nature 328, 726–728.

Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 107, 3588–3593.

Conine, C.C., Moresco, J.J., Gu, W., Shirayama, M., Conte, D., Yates, J.R., and Mello, C.C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell.

Cox, D.N., Chao, a, Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. 12, 3715–3727.

Crick, F.H. (1958). On protein synthesis. Symp. Soc. Exp. Biol. 12, 138–163.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell 31, 79–90.

Djikeng, A., Shi, H., Tschudi, C., and Ullu, E. (2001). RNA interference in Trypanosoma brucei: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. RNA 7, 1522–1530.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science 326, 544–550.

Duchaine, T.F., Wohlschlegel, J. a., Kennedy, S., Bei, Y., Conte, D., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G., et al. (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124, 343–354.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411, 494–498.

Elkayam, E., Kuhn, C.-D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J., and Joshua-Tor, L. (2012). The structure of human argonaute-2 in complex with miR-20a. Cell 150, 100–110.

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P.N., Enright, A.J., and O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature 480, 259–263.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Garcia Silva, M.R., Tosar, J.P., Frugier, M., Pantano, S., Bonilla, B., Esteban, L., Serra, E., Rovira, C., Robello, C., and Cayota, A. (2010). Cloning, characterization and subcellular localization of a Trypanosoma cruzi argonaute protein defining a new subfamily distinctive of trypanosomatids. Gene 466, 26–35.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442, 199–202.

Goh, W.S., Seah, J.W.E., Harrison, E.J., Chen, C., Hammell, C.M., and Hannon, G.J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. Genes Dev. 28, 797–807.

Goriaux, C., Desset, S., Renaud, Y., Vaury, C., and Brasset, E. (2014). Transcriptional properties and splicing of the flamenco piRNA cluster. EMBO Rep. 15, 411–418.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell 106, 23–34.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151, 1488–1500.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science 321, 537–541.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature 465, 1097–1101.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol. 17, 1265–1272.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D. V, Blaser, H., Raz, E., Moens, C.B., et al. (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129, 69–82.

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature 491, 279–283.

Izant, J.G., and Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. Cell 36, 1007–1015.

Izumi, N., Kawaoka, S., Yasuhara, S., Suzuki, Y., Sugano, S., Katsuma, S., and Tomari, Y. (2013). Hsp90 facilitates accurate loading of precursor piRNAs into PIWI proteins. RNA 19, 896–901.

Kamminga, L.M., van Wolfswinkel, J.C., Luteijn, M.J., Kaaij, L.J.T., Bagijn, M.P., Sapetschnig, A., Miska, E.A., Berezikov, E., and Ketting, R.F. (2012). Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans. PLoS Genet. 8, e1002702.

Kasper, D.M., Wang, G., Gardner, K.E., Johnstone, T.G., and Reinke, V. (2014). The C. elegans SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. Dev. Cell 31, 145–158.

Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3' end formation of PIWI-interacting RNAs in vitro. Mol. Cell 43, 1015–1022.

Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature 427, 645–649.

Kennerdell, J.R., and Carthew, R.W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell 95, 1017–1026.

Ketting, R.F., Haverkamp, T.H. a, Van Luenen, H.G. a M., and Plasterk, R.H. a (1999). mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of werner syndrome helicase and RNaseD. Cell 99, 133–141.

Khurana, J.S., Wang, J., Xu, J., Koppetsch, B.S., Thomson, T.C., Nowosielska, A., Li, C., Zamore, P.D., Weng, Z., and Theurkauf, W.E. (2011). Adaptation to P element transposon invasion in Drosophila melanogaster. Cell 147, 1551–1563.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., et al. (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell 138, 1137–1149.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31, 147–157.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science 313, 363–367.

Lee, F., and Yanofsky, C. (1977). Transcription termination at the trp operon attenuators of Escherichia coli and Salmonella typhimurium: RNA secondary structure and regulation of termination. Proc. Natl. Acad. Sci. U. S. A. 74, 4365–4369.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150, 78–87.

Lee, J.T., Strauss, W.M., Dausman, J. a., and Jaenisch, R. (1996). A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell 86, 83–94.

Lin, H., and Spradling, a C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development 124, 2463–2476.

Liu, J., Carmell, M. a, Rivas, F. V, Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science 305, 1437–1441.

Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO J. 31, 3422–3430.

Maison, C., Romeo, K., Bailly, D., Dubarry, M., Quivy, J.-P., and Almouzni, G. (2012). The SUMO protease SENP7 is a critical component to ensure HP1 enrichment at pericentric heterochromatin. Nat. Struct. Mol. Biol. 19, 458–460.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell 136, 656–668.

Malone, C.D., Anderson, A.M., Motl, J.A., Rexer, C.H., and Chalker, D.L. (2005). Germ line transcripts are processed by a Dicer-like protein that is essential for developmentally programmed genome rearrangements of Tetrahymena thermophila. Mol. Cell. Biol. 25, 9151–9164.

Matzke, M.A., Primig, M., Trnovsky, J., and Matzke, A.J. (1989). Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J. 8, 643–649.

Michelitsch, M.D., and Weissman, J.S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. Proc. Natl. Acad. Sci. U. S. A. 97, 11910–11915.

Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The Rhino-Deadlock-Cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. Cell 157, 1364–1379.

Montgomery, T.A., Rim, Y.-S.S., Zhang, C., Dowen, R.H., Phillips, C.M., Fischer, S.E.J., and Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. PLoS Genet. 8, e1002616.

Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J., et al. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature 491, 284–287.

Nowacki, M., Vijayan, V., Zhou, Y., Schotanus, K., Doak, T.G., and Landweber, L.F. (2008). RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature 451, 153–158.

Olivieri, D., Senti, K.-A., Subramanian, S., Sachidanandam, R., and Brennecke, J. (2012). The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila. Mol. Cell 47, 954–969.

Pane, A., Jiang, P., Zhao, D.Y., Singh, M., and Schüpbach, T. (2011). The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. EMBO J. 30, 4601–4615.

Parker, G.S., Eckert, D.M., and Bass, B.L. (2006). RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. RNA 12, 807–818.

Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the C . elegans germline. 1433–1444.

Phillips, C.M., Montgomery, B.E., Breen, P.C., Roovers, E.F., Rim, Y.-S., Ohsumi, T.K., Newman, M.A., van Wolfswinkel, J.C., Ketting, R.F., Ruvkun, G., et al. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. Curr. Biol. 24, 839–844.

Preall, J.B., Czech, B., Guzzardo, P.M., Muerdter, F., and Hannon, G.J. (2012). shutdown is a component of the Drosophila piRNA biogenesis machinery. RNA 18, 1446–1457.

Reinke, V., Gil, I.S., Ward, S., and Kazmer, K. (2004). Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development 131, 311–323.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127, 1193–1207.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev. 21, 1603–1608.

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Dev. Cell 27, 656–663.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150, 65–77.

Shirayama, M., Stanney, W., Gu, W., Seth, M., and Mello, C.C. (2014). The Vasa homolog RDE-12 engages target mRNA and multiple argonaute proteins to promote RNAi in C. elegans. Curr. Biol. 24, 845–851.

Stark, B.C., Kole, R., Bowman, E.J., and Altman, S. (1978). Ribonuclease P: an enzyme with an essential RNA component. Proc. Natl. Acad. Sci. U. S. A. 75, 3717–3721.

Steiner, F. a, Hoogstrate, S.W., Okihara, K.L., Thijssen, K.L., Ketting, R.F., Plasterk, R.H. a, and Sijen, T. (2007). Structural features of small RNA precursors determine Argonaute loading in Caenorhabditis elegans. Nat. Struct. Mol. Biol. 14, 927–933.

Steiner, F.A., Okihara, K.L., Hoogstrate, S.W., Sijen, T., and Ketting, R.F. (2009). RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. Nat. Struct. Mol. Biol. 16, 207–211.

Strome, S., and Wood, W.B. (1982). Immunofluorescence visualization of germ-line-specific cytoplasmic granules in embryos, larvae, and adults of Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 79, 1558–1562.

Swarts, D.C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R.F., Koonin, E. V, Patel, D.J., and van der Oost, J. (2014). The evolutionary journey of Argonaute proteins. Nat. Struct. Mol. Biol. 21, 743–753.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99, 123–132.

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-Box helicase to direct RNAi in C. elegans. Cell 109, 861–871.

Temin, H.M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature 226, 1211–1213.

Thivierge, C., Makil, N., Flamand, M., Vasale, J.J., Mello, C.C., Wohlschlegel, J., Conte, D., and Duchaine, T.F. (2011). Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. Nat. Struct. Mol. Biol. 19, 90–97.

Tijsterman, M., Ketting, R.F., Okihara, K.L., Sijen, T., and Plasterk, R.H.A. (2002). RNA helicase MUT-14-dependent gene silencing triggered in C. elegans by short antisense RNAs. Science 295, 694–697.

Tops, B.B.J., Tabara, H., Sijen, T., Simmer, F., Mello, C.C., Plasterk, R.H. a, and Ketting, R.F. (2005). RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. Nucleic Acids Res. 33, 347–355.

Tsai, H.-Y., Chen, C.-C.G., Conte, D., Moresco, J.J., Chaves, D.A., Mitani, S., Yates, J.R., Tsai, M.-D., and Mello, C.C. (2015). A Ribonuclease Coordinates siRNA Amplification and mRNA Cleavage

during RNAi. Cell 160, 407–419.

Vagin, V. V, Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science 313, 320–324.

Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U. S. A. 107, 3582–3587.

Vastenhouw, N.L., Fischer, S.E.J., Robert, V.J.P., Thijssen, K.L., Fraser, A.G., Kamath, R.S., Ahringer, J., and Plasterk, R.H. a. (2003). A Genome-Wide Screen Identifies 27 Genes Involved in Transposon Silencing in C. elegans. Curr. Biol. 13, 1311–1316.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297, 1833–1837.

Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. Curr. Biol. 18, 861–867.

Wargelius, A., Ellingsen, S., and Fjose, A. (1999). Double-stranded RNA induces specific developmental defects in zebrafish embryos. Biochem. Biophys. Res. Commun. 263, 156–161.

Wassenegger, M., Heimes, S., Riedel, L., and Sänger, H.L. (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell 76, 567–576.

Waterhouse, P.M., Graham, M.W., and Wang, M.B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. U. S. A. 95, 13959–13964.

Watson, J., and Crick, F.H.F. (1953). Molecular structure of nucleic acids. Nature 171, 737–738.

Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013). Protection of Germline Gene Expression by C. elegans Argonaute CSR-1. Dev. Cell.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R. a, Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E. a (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. 28, 783–796.

Van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell 139, 135–148.

Yang, H., Zhang, Y., Vallandingham, J., Li, H., Florens, L., and Mak, H.Y. (2012). The RDE-10/RDE-11 complex triggers RNAi-induced mRNA degradation by association with target mRNA in C. elegans. Genes Dev. 26, 846–856.

Yang, H., Vallandingham, J., Shiu, P., Li, H., Hunter, C.P., and Mak, H.Y. (2014). The DEAD box helicase RDE-12 promotes amplification of RNAi in cytoplasmic foci in C. Elegans. Curr. Biol. 24, 832–838.

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.-C.G., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127, 747–757.

Zhang, C., Montgomery, T. a, Gabel, H.W., Fischer, S.E.J., Phillips, C.M., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 108, 1201–1208.

Zhang, C., Montgomery, T.A., Fischer, S.E.J., Garcia, S.M.D.A., Riedel, C.G., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2012). The Caenorhabditis elegans RDE-10/RDE-11 complex regulates RNAi by promoting secondary siRNA amplification. Curr. Biol. 22, 881–890.

Zhang, Z., Wang, J., Schultz, N., Zhang, F., Parhad, S.S., Tu, S., Vreven, T., Zamore, P.D., Weng, Z., and Theurkauf, W.E. (2014). The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. Cell 157, 1353–1363.

Chapter 2

PID-1 is a novel factor that operates during 21U-RNA biogenesis in *Caenorhabditis elegans*.

Bruno F.M. de Albuquerque, Maartje J. Luteijn, Ricardo Rodrigues, Petra van Bergeijk, Selma Waaijers, Lucas J.T. Kaaij, Holger Klein, Mike Boxem and René F. Ketting

SUMMARY

The Piwi-piRNA pathway represents a small RNA-based mechanism responsible for the recognition and silencing of invading DNA. Biogenesis of piRNAs (21U-RNAs) is poorly understood. In *Caenorhabditis elegans*, the piRNA-binding Argonaute protein PRG-1 is the only known player acting downstream from precursor transcription. From a screen aimed at the isolation of piRNA-induced silencing-defective (Pid) mutations, we identified, among known Piwi pathway components, PID-1 as a novel player. PID-1 is a mostly cytoplasmic, germline-specific factor essential for 21U-RNA biogenesis, affecting an early step in the processing or transport of 21U precursor transcripts. We also show that maternal 21U-RNAs are essential to initiate silencing.

INTRODUCTION

One of the challenges faced by germ cells is the faithful inheritance of their genomes over generations. An important aspect of this process is to ensure that potentially harmful DNA sequences are not expressed and/or do not replicate. Such sequences include transposons and retroviruses and, when left unchecked, create havoc, leading to fertility problems. One of the mechanisms controlling this line of defense is the Piwi–piRNA system (Ghildiyal and Zamore, 2009; Luteijn and Ketting, 2013; Malone and Hannon, 2009). The Piwi–piRNA pathway is a small RNA-based mechanism. At its core, Argonaute proteins from the Piwi subfamily use small RNA cofactors called piRNAs to recognize and silence transcripts derived from repetitive elements. Silencing often involves the amplification of a small RNA population and can occur at both the posttranscriptional and transcriptional levels. Post-transcriptional silencing involves cleavage of targeted mRNA molecules by the RNase H activity of Piwi proteins, while transcriptional silencing likely involves the recognition of a nascent transcript and is accompanied by the acquisition of repressive chromatin marks at the targeted locus (Ashe et al., 2012; Buckley et al., 2012; Huang et al., 2013; Luteijn et al., 2012; Rozhkov et al., 2013; Shirayama et al., 2012; Sienski et al., 2012; Le Thomas et al., 2013). This heterochromatin-like state can spread and lead to the silencing of genes that are located close to a Piwi target (Sienski et al., 2012). Despite the fact that a number of proteins have been linked to these activities (Ghildiyal and Zamore, 2009; Luteijn and Ketting, 2013; Malone and Hannon, 2009), the exact mechanisms of this transcriptional silencing mechanism remain unclear. Although the core principles of Piwi-driven mechanisms appear to be strongly conserved, the precise mechanisms through which Piwi proteins act and piRNAs are produced are remarkably flexible between species. In the nematode *Caenorhabditis elegans*, the Piwi protein PRG-1 is guided to its targets by piRNAs (named 21U-RNAs in C. elegans) (Batista et al., 2008; Das et al., 2008; Ruby et al., 2006; Wang and Reinke, 2008) in a process that tolerates significant mismatches (Bagijn et al., 2012; Lee et al., 2012). Combined with the fact that; 30,000 21U-RNAs have been identified (Gu et al., 2012), 21U-RNAs can identify almost any foreign DNA sequence. Following target recognition by PRG-1, a second class of small RNAs (22G-RNAs) is produced, using the target transcript as a template (Das et al., 2008). In order to produce 22G-RNAs, presumably one of the RNA-dependent RNA polymerase (RdRP) enzymes of *C. elegans* is required as well as some genes

known as mutator genes, like *mut-7* and *rde-3* (Bagijn et al., 2012). The 22G-RNAs are bound by a different set of Argonaute proteins, including the nuclear protein HRDE-1/WAGO-9 (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). The final outcome is reduced transcription and tri-methylation of Lys9 of histone H3 (H3K9me3) of PRG-1 targeted genes (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Intriguingly, this transcriptional silencing response can acquire a PRG-1-independent, self-maintaining state. Under these conditions, HRDE-1, RDE-3, and MUT-7 are still required (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). The least-understood step in this PRG-1-mediated mechanism is how 21U-RNAs are made. In fact, we only know that 21U-RNA genes are marked by the presence of an upstream consensus motif (Ruby et al., 2006) of one transcription factor that plays a role in their transcription (Cecere et al., 2012) and that they are made as 5'-capped precursors that are at least 26 nucleotides (nt) in length (Gu et al., 2012). From these precursors, the two most 5' bases are removed, likely followed by binding to PRG-1 and 3' end trimming (Kawaoka et al., 2011). We do not know any of the factors acting during precursor processing. Here, we describe a forward mutagenesis screen aimed at the isolation of mutants that are piRNA-induced silencing-defective (Pid). This screen led to the identification of alleles of known PRG-1 pathway components but also revealed a novel protein, PID-1, that plays an important role in 21U-RNA biogenesis downstream from their transcription. We also describe genetic experiments revealing intriguing maternal and paternal effects of the PRG-1 pathway.

RESULTS AND DISCUSSION

Isolation of mutants in the 21U pathway.

In order to identify novel genes acting in the PRG-1- mediated silencing pathway, we made use of a strain expressing a GFP-Histone-2B fusion gene that carries a specific 21UR1-RNA recognition site in its 3' untranslated region (UTR) (21U sensor) (Bagijn et al., 2012). We previously described how loss of HENN-1, the enzyme catalyzing the 3' end methylation of 21U-RNAs, leads to partial de-silencing of this transgene (Kamminga, 2011). We also demonstrated that this partially activated state is subject to spontaneous re-silencing, after which this sensor construct is no longer responsive to removal of PRG-1 (Luteijn et al., 2012).

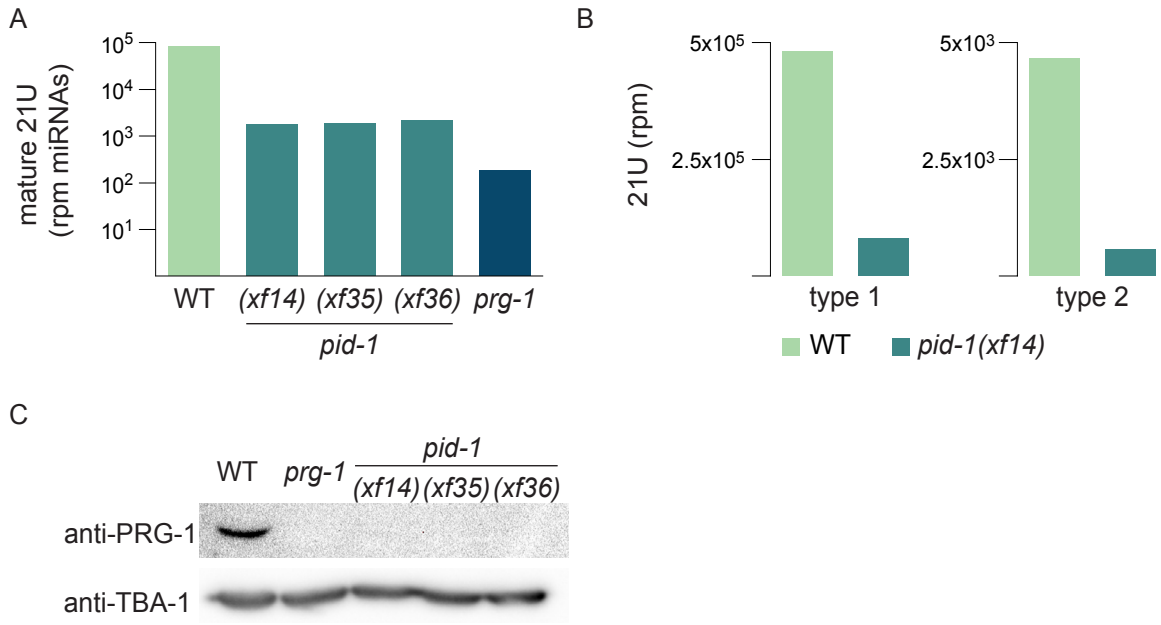


Figure 1. *pid-1* affects 21U-RNAs and PRG-1. (A) Bar diagram displaying normalized mature 21U read counts in strains with the indicated genotypes. Normalization was done to total miRNA read counts. Mature 21U reads were defined as those reads starting at the annotated mature 59 end. **(B)** Bar diagram displaying type 1 and type 2 21U-RNA abundance in the indicated strains. Read counts reflect 21U counts obtained from $NaIO_4$-oxidized small RNA populations normalized to total mapped non-rRNA reads. **(C)** Western blotting analysis of extracts from wild-type, *prg-1*, and *pid-1* mutant animals (young adults). Tubulin was used as a loading control.

In order to identify genes that act at any stage of the piwi pathway, we mutagenized a population of animals carrying a partially activated 21U sensor in a *henn-1* mutant background with EMS and scored for full activation of the 21U sensor (Supplemental Fig. S1A). Genome-wide re-sequencing of retrieved mutant strains identified premature stop mutations in *mut-7* and *rde-3* (Supplemental Fig. S1B). We also identified an interesting allele (*xf25*) of *hrde-1* carrying a non-synonymous mutation affecting the last amino acid of HRDE-1 (A1032T). Given the important role for the most C-terminal residue of Argonaute proteins in making the binding pocket for the 5' phosphate of the small RNA cofactor (Patel et al., 2006), we reasoned that this mutation might disrupt HRDE-1 activity. Indeed, a deletion allele of *hrde-1*, *tm1200*, was unable to complement *xf25* (data not shown), showing that HRDE-1(A1032T) is nonfunctional. In summary, our screen effectively identifies genuine components of the *C. elegans* piwi pathway.

PID-1, a novel, general 21U-RNA biogenesis factor.

Most of the alleles identified in our screen displayed normal levels of 21U-RNAs,

as assessed by Northern blotting (Supplemental Fig. S1C) except for xf14. The locus affected by *xf14* was named *pid-1*. To probe the generality of the apparent 21U accumulation defect, we analyzed small RNAs from both wild-type and *pid-1(xf14)* young adult animals that also carry the 21U sensor through deep sequencing (Supplemental Table S1). Both libraries have a very similar size and display similar microRNA (miRNA) profiles (Supplemental Table S1; Supplemental Fig. S2A), indicating that the libraries can be well compared. As expected, in wild-type animals, we could detect small RNAs mapping to the 21U sensor (Supplemental Fig. S2B). These small RNAs display all characteristics of 22G-RNAs (Supplemental Fig. S3). While in *pid-1(xf14)* mutants these small RNAs still display all 22G characteristics, their levels are strongly reduced (Supplemental Fig. S2B). We next checked mature 21U-RNA levels, normalizing to total miRNA counts. We found that, overall, 21U-RNA levels were 10-fold reduced in *pid-1(xf14)* mutants (Fig. 1A) and that most 21U-RNA loci responded similarly to *pid-1(xf14)* (Supplemental Fig. S2A). The loss of 21URNAs is slightly less than observed in *prg-1(pk2298)* mutants (Fig. 1A), suggesting that 21U-RNA biogenesis can proceed partially in the absence of PID-1. The 59 nucleotide composition of the remaining 21U-RNAs in the pid-1 mutants was still heavily biased toward uracil (Supplemental Fig. S2C). Finally, we asked whether PID-1 equally affects both type 1 and type 2 21U-RNAs. Most 21U-encoding loci are characterized by a specific sequence motif upstream of the transcription start site (Ruby et al., 2006), and a later study showed that this motif may be a transcription factor-binding site (Cecere et al., 2012). However, it was shown recently that many loci exist that do not have this sequence motif and still produce 21U-RNAs. These two different classes have been referred to as type 1 and type 2 21U-RNAs, respectively (Gu et al., 2012). We found that *pid-1(xf14)* leads to a strong decrease in type 1 as well as type 2 21U-RNAs (Fig. 1B), indicating that PID-1 does not affect recognition of the upstream 21U sequence motif. We conclude that PID-1 is a factor involved in global 21U-RNA biogenesis.

PID-1 affects PRG-1 levels and inhibits RNAi.

The only gene thus far known to be required for 21U-RNA accumulation is *prg-1*. Hence, PID-1 could be required for transcription of *prg-1*. In order to test this, we performed RT-qPCR on wild-type and *pid-1* mutant animals (Supplemental Fig. S4A). This revealed that transcript levels of *prg-1* are unaffected by *pid-1*. In

contrast, PRG-1 protein levels were strongly reduced upon mutation of *pid-1* (Fig. 1C), indicating a post- transcriptional effect of *pid-1* on PRG-1. Although we formally cannot exclude a role for PID-1 in PRG-1 translation, we interpret this result as instability of PRG-1 in the absence of 21U-RNAs, much like Drosophila Piwi is unstable without piRNAs, and Argonaute proteins are destabilized by loss of miRNAs (Olivieri et al., 2010; Smibert et al., 2013). Since many PRG-1 pathway components affect RNAi triggered by exogenous dsRNA, we also probed the RNAi sensitivity of *pid-1* mutants. Using pos-1 RNAi by feeding, we found that *pid-1* mutants, like *prg-1* mutants, are slightly hypersensitive (Supplemental Fig. S4B). Most likely, this reflects competition for components shared by the exo-RNAi machinery and the *prg-1* pathway, such as MUT-7 and RDE-3.

Identification of *pid-1* as F18A1.8

To identify the *pid-1* gene, we first crossed the *pid-1(xf14)* allele into a *henn-1* wild-type background. This showed that the loss of 21U-RNAs in *pid-1(xf14)* animals

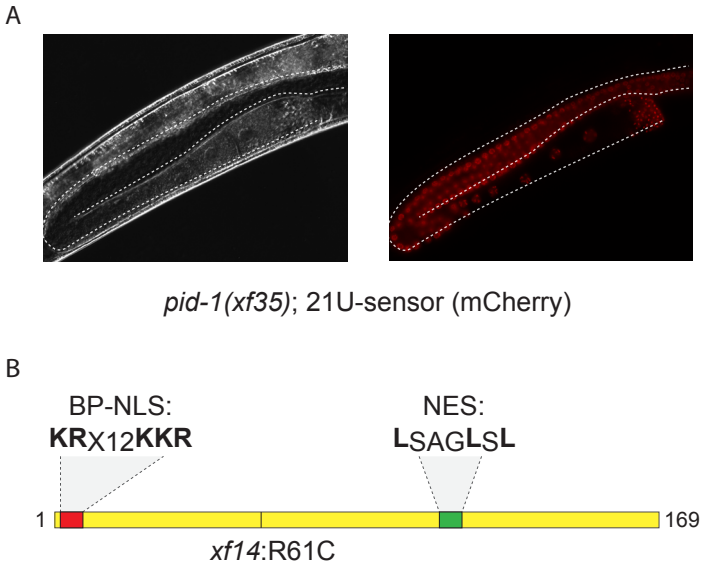


Figure 2. Identity of *pid-1*. (A) A typical *pid-1(xf35)* mutant animal carrying a mCherry-21U sensor transgene that has been crossed in from a *mut-7(pk204)* mutant background. The depicted animal is homozygous wild type for *mut-7*. **(B)** Schematic of the PID-1 protein.

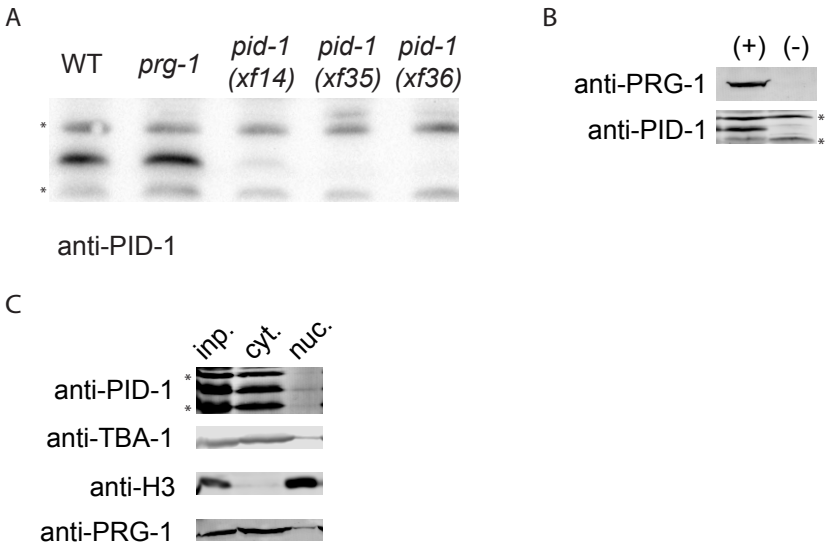


Figure 3. PID-1 is a mostly cytoplasmic, germline-expressed protein. (A) Western blot for PID-1 in the indicated backgrounds. The asterisks mark nonspecific signals from the custom-raised antiserum. The non-specific signals were also used as a loading control. **(B)** Western blot detecting PID-1 and PRG-1 protein in *glp-4(bn2ts)* mutant animals grown at permissive (with germline) and restrictive (without germline) temperatures. The asterisks mark nonspecific signals from the custom-raised antiserum. **(C)** Western blot analysis of a subcellular fractionation experiment. TBA-1 was used to validate cytoplasmic fractions (cyt), and Histone H3 was used to validate nuclear fractions (nuc).

is *henn-1*-independent (Supplemental Fig. S5A). Genetic mapping revealed that *pid-1* is genetically closely linked to the 21U sensor transgene on chromosome II (data not shown), and genome-wide resequencing revealed three non-silent mutations in the area where *pid-1* was mapped; these three mutations affect F18A1.8, ZK1320.5, and F37H8.3. We then selected recombinants that separated the 21U sensor from these three mutations. Only recombinants that picked up a wild-type copy of F18A1.8 produced 21URNAs (Supplemental Fig. S5A), suggesting that F18A1.8 is *pid-1*. To test this, we generated two additional alleles (*xf35* and *xf36*) of F18A1.8 using CRISPR technology and probed the effect of these alleles on the activity of our 21U sensor and on 21U-RNA biogenesis. First, *xf35* and *xf36* do not complement *pid-1(xf14)* (Supplemental Fig. S5B; data not shown). Second, *pid-1(xf35)* was unable to silence a 21U cherry sensor transgene coming from a *mut-7(pk204)* mutant background (Fig. 2A). Third, both *xf35* and *xf36* have a similar effect on mature 21U-RNA levels as *pid-1(xf14)*, as revealed by deep sequencing (Fig. 1A). Finally, like in pid-1(xf14) mutants, PRG-1 levels are strongly reduced in *xf35* and *xf36* mutant extracts (Fig. 1C). Taken together, these data prove that *pid-1* corresponds to F18A1.8. PID-1 is predicted to be a protein of 169 amino acids,

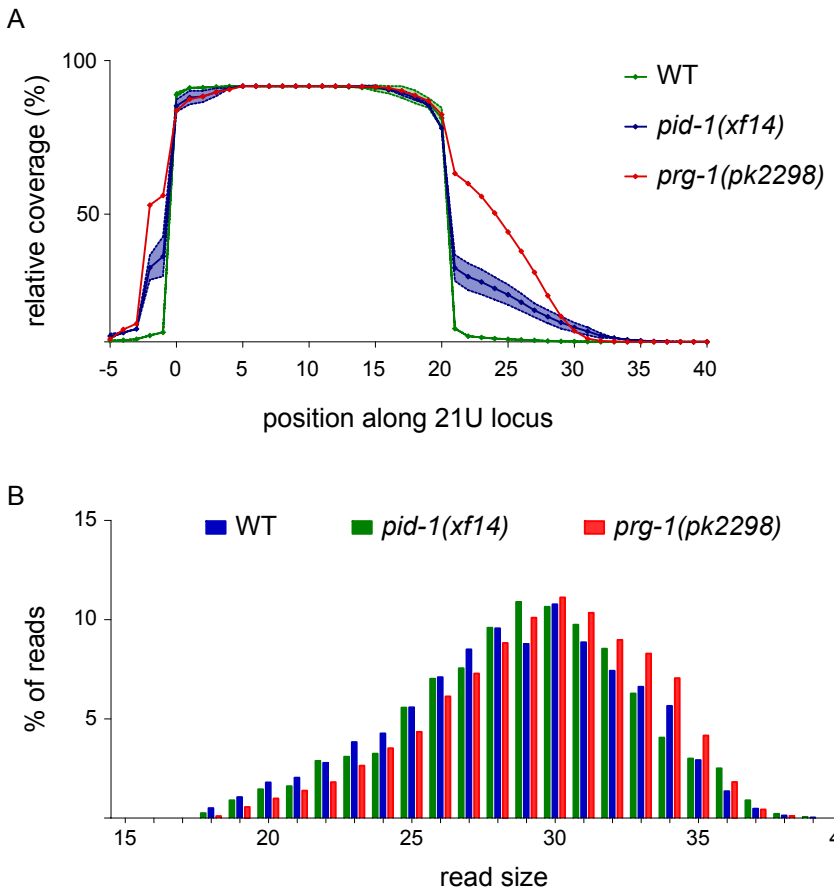


Figure 4. PID-1 acts during 21U precursor processing. (A) Relative coverage of individual bases of 21U loci by obtained cDNA reads from *prg-1*, *pid-1*, and wild-type libraries made from TAP-treated small RNA. The annotated 59 ends of the analyzed loci are at position 0. The dotted lines reflect the standard deviation from the two wild-type samples and the three pid-1 alleles. **(B)** A length distribution plot for 21U-RNA precursor reads obtained from the CIP-TAP-treated libraries.

and *pid-1(xf14)* results in an arginine-to-cysteine change at position 61 (R61C) (Fig. 2B). This mutation somehow affects PID-1 stability, since Western blotting reveals that PID-1 expression is strongly reduced in *pid-1(xf14)* mutant extracts, while the *xf35* and *xf36* alleles do not produce detectable PID-1 protein at all (Fig. 3A). Next, we tested whether PID-1, like PRG-1, is expressed specifically in the germline by performing Western blot analysis on mutants with a temperature-sensitive allele of *glp-4*, resulting in a lack of germline tissue in animals grown at 25°C. This revealed that PID-1 is expressed predominantly in the germline. This is consistent with published data showing that *pid-1* transcripts are selectively enriched in the primordial germ cells of C. elegans (Spencer et al., 2011). No obvious domain structure is detected through regular BLAST analysis. However, at the N terminus, a bipartite nuclear localization signal may be present (KREFSHITLASTPFKKR), and the protein may

contain a nuclear export signal (LSAGLSL) as well (Fig. 2B), suggesting that PID-1 may be cycling between the nucleus and cytoplasm. Subcellular fractionation experiments suggest that the majority of PID-1 is cytoplasmic (Fig. 3C). However, given that the PID-1 signal is retained better in the nuclear fraction than the non-PID-1-specific signals, it is possible that a small part of the PID-1 pool may be associated with or be inside the nucleus (Fig. 3C).

PID-1 does not affect 21U-RNA methylation

To further probe at which level PID-1 affects 21U-RNA biosynthesis, we NaIO4-oxidized small RNA. This procedure enriches for small RNA species that carry a 2-O-methyl modification at their 3' end, such as mature 21URNAs in C. elegans (Ruby et al., 2006), and it is believed that this methylation step finalizes 21U-RNA biogenesis. To assess the effect of the oxidation, we compared the ratio of 21U with miRNA reads in the tobacco alkaline phosphatase (TAP)-treated and the oxidized RNA-derived libraries. Using this analysis, we were unable to detect a significant difference between the wild type and *pid-1* mutants (Supplemental Table S1), indicating that the remaining 21U-RNAs in *pid-1(xf14)* mutants are still 2-O-methylated. Furthermore, the global requirement of *pid-1(xf14)* for 21U-RNA accumulation is also observed in the libraries made from oxidized RNA (Supplemental Fig. S6A). We also checked whether the step right before methylation is affected by PID-1. This step involves the trimming of the 21U3' end, presumably by an exonuclease. To assess this, we profiled the length distribution of those 21U reads from the TAP-treated libraries whose 5' ends match precisely onto an annotated 5' end of a mature 21U-RNA, indicating that the 5' end has been processed (see below). This revealed no differences between the wild-type and mutant 21U-RNAs (Supplemental Fig. S6B). We conclude that PID-1 does not affect the 3' end processing of 21U-RNAs.

PID-1 acts during 21U precursor processing

We also probed our libraries for potential 21U-RNA precursor transcripts. It has been suggested before that these precursors mostly start 2 nt upstream of the 5' end of the mature 21U-RNA, are 5'-capped, and are at least 26 nt in length (Gu et al., 2012). These species should be represented in our TAP-treated libraries, as

TAP has decapping activity. Hence, we analyzed all reads that come from unique 21U loci (Supplemental Table S2) and start within 5 nt of the annotated mature 5' end. In order to prevent strong biases by heavily expressed individual 21U loci, we restricted our analysis to those 21U loci that individually produce <1% of the 21U reads. We then plotted read coverage over these 21U loci, starting at position -5 and extending until position +40, with the start of the annotated mature 21U-RNA at position 0. First, such analysis in two replicate samples of wild-type animals reveals that this approach yields very reproducible patterns (Supplemental Fig. S6C).We then performed the same analysis on prg-1(pk2298) small RNA-derived cDNA libraries. This revealed that in the absence of PRG-1 , ~80% of the reads derived from 21U loci correspond to precursor transcripts, starting mostly at position -2 and extending to ~35 nt in length (Fig. 4A). We then analyzed the effects of the three *pid-1* alleles (Fig. 4A; Supplemental Fig. S6C), revealing a clear accumulation of 21U-RNA precursors in animals lacking PID-1. Interestingly, the effect of *pid-1* is somewhat less than that of *prg-1*. Given the fact that *pid-1* mutants still produce low levels of 21U-RNAs while *prg-1* mutants almost lack them completely, the simplest interpretation of these data is that in *pid-1* mutants, 21U precursors can still be partially processed into mature 21URNAs. Finally, we profiled the length and locus coverage of 21U precursor transcripts in the diverse wild-type and mutant backgrounds using a protocol that selects for capped transcripts (Gu et al., 2012). This experiment did not reveal significant differences in the 21U precursor structure between any of the tested strains, suggesting that PID-1 does not affect the structure of 21U precursor transcripts (Fig. 4B).

Genetic requirements for *pid-1* and *prg-1* during transgene (re)silencing.

Finally, in order to better understand the global role of 21U-RNAs during the process of transgene silencing, we assessed how active transgenes from 21U-defective backgrounds react to the re-establishment of the 21U pathway. For that, we crossed *pid-1(xf14);21U sensor* males into either wild-type or *pid-1* mutant hermaphrodites. As expected, offspring of these males mated with *pid-1*- defective hermaphrodites show strong transgene activity. Interestingly, however, when mated to wild-type hermaphrodites, these males sire offspring with active transgenes, albeit the fluorescence intensity is notably lower (Table 1). We observed similar transgene activity in offspring from *prg-1(pk2298);21U sensor* males mated with

Table 1. Parental effects in the 21U pathway.

Genotype	Mother	Father	Fluorescence
pid-1(xf14)	*pid-1(xf14)*	*pid-1(xf14); sensor*	++
prg-1(pk2298)	*prg-1(pk2298)*	*prg-1(pk2298); sensor*	++
pid-1(xf14)/+ (F1)	+/+	*pid-1(xf14); sensor*	+
pid-1(xf14)/+ (F2)	Selfing of above F1		-
prg-1(pk2298)/+ (F1)	+/+	*prg-1(pk2298); sensor*	+
prg-1(pk2298)/+ (F2)	Selfing of above F1		-
mut-7(pk204)/+ (F1)	+/+	*mut-7(pk204); sensor*	-
mut-7(pk204)/+; pid-1(xf14)/+ (F1)	*pid-1(xf14)*	*mut-7(pk204); sensor*	++

All analyzed animals were offspring from a 21U sensor-carrying male crossed with a non-transgenic hermaphrodite. Genotypes of the parents are indicated. Fluorescence intensity was scored under a dissection microscope. (++) Detectable at 120X (+) detectable at 480X; (-) not detectable at 480X.

wild-type hermaphrodites. Full silencing of the transgene is only observed in the next generation (Table 1). This shows that while 21U-RNAs start to silence a novel target within one generation, full silencing can be achieved only upon joint passage through the germline. Possibly, 22GRNA levels need to accumulate in order to establish full silencing. Next, we crossed an active 21U sensor from *mut-7(pk204)* males into both wild-type and *pid-1* mutant hermaphrodites and analyzed transgene expression in the F1. This revealed a very strong maternal effect of 21U-RNAs: The wild-type hermaphrodites produced fully silenced offspring, while the progeny of the *pid-1* mutant mothers could not initiate detectable silencing even though they were heterozygous wild type for *pid-1*. This result highlights at least two properties of the 21U pathway. First, the maternally provided 21U-RNA pool is essential to establish re-silencing of a *mut-7(pk204)*-activated transgene, suggesting that the initiation of silencing by 21U-RNAs happens during embryonic development using maternally provided 21U-RNAs. Second, a *mut-7*-activated transgene coming in via the sperm is easier to re-silence by 21URNAs than a *pid-1* or a *prg-1*-activated transgene. We conclude that some silencing information must be present in *mut-7*-derived sperm but is lacking in *prg-1*-defective sperm and that this information can be efficiently used by the maternal 21U-RNA pool.

CONCLUSIONS

We identified a novel gene, *pid-1*, that plays an important role during 21U-RNA

biogenesis in C. elegans. PID-1 is a mostly cytoplasmic protein that likely acts during the processing of 21U-RNAs downstream from their transcription. The mature 21U-RNAs that are still made in *pid-1* mutants are indistinguishable from wild-type 21URNAs, and, relative to these mature 21U-RNAs, 21U-RNA precursors accumulate upon loss of PID-1. Furthermore, the structure of 21U precursor transcripts is not affected by PID-1. Given these results as well as the potential nuclear import and export signals on PID-1, we suggest that PID-1 may somehow guide 21U precursors to a site where they are further processed into mature 21U-RNAs. We note that in other species, like mice or Drosophila, PID-1 orthologs cannot be identified through BLAST analysis. This may be related to the fact that, overall, the 21U-RNA pathway in C. elegans deviates significantly from piRNA pathways in other animals. However, given the still poor characterization of piRNA biogenesis in general, proteins with similar functions but with diverged sequences may operate in mammalian or fly piRNA biogenesis. We found that pid-1 as well as prg-1 partially inhibit RNAi triggered by exogenous dsRNA .Much like what is described for other RNAi-hypersensitive mutants, we attribute this to the fact that the prg-1 pathway included proteins shared with the exo-RNAi pathway (Duchaine et al., 2006; Pavelec et al., 2009), and thus ongoing prg-1 activity in C. elegans puts a significant claim on those shared resources. On the genetic level, we showed that the initiation of transgene silencing by 21U-RNAs depends on maternally provided 21U-RNAs. These findings are consistent with the maternal contributions of piRNA populations observed in other species, like Drosophila (Brennecke et al., 2008) and zebrafish (Houwing et al., 2007), and highlight the important role for maternally provided small RNA populations in order to establish silencing in offspring. In addition, we found that the genetic background of the sperm can have an effect on the silencing of a locus. Considered together, parental influences play important roles during the establishment of the epigenetic state of an allele through the 21U pathway in C. elegans.

ACKNOWLEDGMENTS

We thank Tina Han for valuable assistance during library preparation and sequencing, and Joern Toedling, Nastasja Kreim, and Ies Nijdam for bioinformatics support. We thank Miguel Almeida for technical assistance during the genetic mapping experiments. This work was supported by grants to R.F.K. (European

Research Council Starting Grant from the Ideas Program of the European Union Seventh Framework Program, grant no. 202819, and two grants from the Netherlands Organization for Scientific Research [NWO]: ECHO 700.57.006 and Vici 724.011.001) and B.F.M.d.A. (Fundação para a Ciência e Tecnologia [FCT], SFRH/BD/51001/ 2010).

MATERIALS AND METHODS

C. elegans strains and culture

C. elegans was cultured on NGM plates, with *Escherichia coli* OP50 as a food source. Alleles that were used in this study are listed in the Supplemental Material.

CAS9-mediated disruption of F18A1.8

To generate additional mutations in F18A1.8, we targeted a site in the first exon (GGAGTTTTCGCATATTACTT) with CRISPR/Cas9 as previously described (Waaijers et al. 2013). The target sequence was cloned into the U6::sgRNA vector pMB70 by annealing oligonucleotides F18_sgRNA_F: 5'-aattGGAGTTTTCGCATATTACTT and F18_sgRNA_R: 5'-aaacAAGTAATATGCGAAAACTCC, and ligating the resulting linker into BsaI digested pMB70. We injected 15 N2 animals with a mixture containing 5 ng/µl Pmyo-3::mCherry (pCFJ104, Addgene #19328), 50 ng/µl F18A1.8 sgRNA and 50 ng/µl Phsp-16.48::Cas9 using standard *C. elegans* microinjection procedures. To induce expression from the hsp-16.48 promoter, injected animals were heat shocked for 1 h at 34 °C on agar plates floating in a water bath, 30 min after injection. From 42 transgenic F1 animals expressing mCherry, we PCR amplified a region surrounding the target site using primers F18A1.8_F: 5'-GAATCGGACGAATCAGGAA and F18A1.8_R: 5'-GTGGTGAGCCAGTTCCATAA, and sequenced the resulting amplicons. Three F1 animals were heterozygous for a short deletion creating a frame-shift and truncating early stop. We established homozygous mutant lines from the two F1 animals that were fertile by isolating single F2 animals and determining their genotype by sequence analysis. The sequences of *pid-1* at the targeted site, and that of the two derived alleles are:

pid-1: TTCGCATATTACTTTGGCATCCACACCATTTAAA

pid-1(xf35): TTCGCATATTTGCATATTTGGCATCCACACCATTTAAA
pid-1(xf36): TTCGCATATTCGCATATTTTGGCATCCACACCATTTAAA

Microscopy

The screen for Pid mutants was performed using a Zeiss M2Bio dissecting microscope. Fluorescence images were taken on a Leica DM6000 B upright microscope with 4003 amplification. Silencing onset of active transgenes was analyzed using a Leica M165 FC stereo microscope with a 13 objective (1203 amplification) or a 43 objective (4803 amplification).

RNAi

Sensitivity to RNAi was tested by feeding *E. coli* expressing dsRNA against *pos-1* (Ketting et al. 1999). Animals were placed on RNAi food as L4 larvae and allowed to lay eggs for 2 d, after which survival was scored. Each strain was tested between four and six times. Data analysis Detailed information on data analysis is provided in the Supplemental Material.

PID-1 antibody

Custom, affinity-purified antibodies against PID-1 were ordered from SDIX, using the whole protein sequence of PID-1 as an antigen. The antibody (animal no. Q5941) was used in a 1:200 dilution on Western blots.

Nucleo–cytoplasmic fractionation

Synchronized adult worms were washed with extract buffer (20 mM MOPS at pH 7.5, 40mMNaCl, 90mMKCl, 2mMEDTA, 0.5 mM EGTA, 10%glycerol, 2mMDTT, proteinase inhibitors [Roche Complete ULTRA]) and re-suspended to a final volume of 4 mL. The suspension was then dripped into liquid nitrogen, and the frozen pellets were ground to a fine powder. The powder was transferred to a glass douncer, thawed on ice, and sheared (30 strokes, piston B). The extracts were then centrifuged twice at 200g to remove debris. The supernatant (crude extract) was centrifuged at 2000g to separate the nuclear fraction from the cytoplasmic fraction. Nuclear fractions were

washed twice with extract buffer, whereas the cytoplasmic fraction was centrifuged at 21,000g to remove any remaining debris.

Small RNA library preparation

Wild-type, *prg-1*, *pid-1(xf14);21U sensor*, *pid-1(xf35)* and *pid-1(xf36)* libraries (TAP and/or CIP-TAP) were prepared as follows. RNA was extracted from a synchronized adult population using Trizol (Sigma- Aldrich). Samples were then enriched for small RNAs using the Mirvana kit from Life Technologies. For CIP treated samples 5 ug of small RNAs were treated with 50 U of calf intestine phosphatase (CIP, NEB) at 37C for 1 hour. [reference to gu2012], followed by Trizol extraction and TAP treatment. One µg of enriched small RNA was size-selected between 15- to 40-nt on 15% TBE-urea gel after treating with 10 U of tobacco acid phosphatase (Epicenter) at 37°C for 2 h to digest 5' tri- and di-phosphates to mono-phosphates. Small RNA libraries were prepared using NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England BioLabs) according to the manufacture's instruction. In brief, small RNA was first ligated to the 3' adapter and then the 5' adapter, both for 1 h at 25°C. Adapter-ligated RNA was reverse-transcribed and PCR-amplified for 15 cycles using NEBNext index primers #13-24. The PCR-amplified cDNA construct was purified using AMPure XP beads (Beckman Coulter). Size selection of the small RNA library was done on LabChip XT instrument (Perkin Elmer) using DNA 300 assay kit. Only the fraction containing 135-161 bp was pooled in equal molar ratio. The resulting 2 nM pool was denatured with 5% PhiX spike-in and sequenced as single-read on HiSeq 2500 (Illumina) in rapid mode for 50 cycles using on-board cluster generation. After de-multiplexing, on average 22 million passing filter reads were obtained per sample.

Sequencing data are available at Gene Expression Omnibus, entry GSE55309.

Data processing

Adapter Removal and Read Mapping

Remaining fragments of sequencing adapters on the 3' ends of reads were removed using the tool FLEXBAR version 2.4 (Dodt et al. 2012).

Subsequently a custom reference genome index was generated based on Ensembl Wbcel215 with an additional sequence for the 21U-sensor construct.

The small RNA reads were processed using a customized pipeline based on the Bioconductor package girafe (Toedling et al. 2010). The pipeline includes a read mapping step with bowtie, version 1.0.0 (Langmead et al. 2009), using the parameters -e 50 -a --best --strata -m 20 --nomaqround --tryhard.

Read counts in annotated loci

The girafe pipeline (Toedling et al. 2010) reported reads mapping to known small RNA loci based on Ensembl WBcel215 for known miRNA and 21U RNA loci. A read was counted as mapping to a known locus if it was mapped completely within the annotated locus extended by two nucleotides in both directions in case of miRNAs, and extended by two nucleotides on the 5' end and ten nucleotides on the 3' end in case of 21U RNA loci.

Nucleotide composition

To visualize the nucleotide composition of the reads mapping to the 21U-sensor construct in the different samples, the respective read sequences were extracted from the SAM file and cut after 18 nucleotides. Reads mapping to the second strand were reverse-complemented. Subsequently sequence logos were generated using the WebLogo (Crooks et al. 2004).

Coverage on 21U sensor

Read coverage on the 21U sensor construct was normalized using the number of unique reads mapping to known miRNA loci (WBcel215) in the respective sample. A pseudocount of one was added to the normalized coverage, the log2 of which then was plotted along the construct sequence using custom R scripts.

21U filtering

In order to analyze nucleotide coverage through the 21U locus, all annotated 21U loci (ws215) were filtered for the ones that are at least 5 nucleotides apart (Table S2). All reads mapped were then filtered for the ones mapping to unique 21U loci and align with the annotated 5' end at position zero.

For 21U type analysis, all 21nt long reads mapped were filtered for the ones starting at annotated 5'. Annotated type2 21U loci were retrieved from ws215 whereas type2 loci were retrieved from Gu el al 2012.

RT-qPCR

Worm cDNA was synthetized from total RNA from synchronized adults using ProtoScript M-MuLV First Strand cDNA synthesis Kit (NEB) following the manufacture instructions for Oligo d(T) priming . prg-1 and tbb-2 amplicons were amplified in a Viia7 Real Time PCR System using the following primers:

prg-1_forward: CATAAGGAACTACGATCGTGC;

prg-1_reverse: CGCGTTGTGGATTGTTCT;

tbb-2_forward: CAATGTGCTTGACGTGATCC;

tbb-2_reverse: AGGTTGGGTTGGTGAGTTTG.

Northern **blotting**

Small RNA was isolated from total RNA with the Mirvana kit (LifeTechnologies). Northern blotting was done as described previously (Kamminga et al. 2012). 20ug of small RNA was loaded on a 12% polyacrylamide gel and blotted according to standard procedures. RNA was UV-cross-linked to the membrane before further handling. Probes were labeled with γ-ATP and PNK. Probe sequences:

21U-R1: 5'-GCACGGTTAACGTACGTACCA

let-7: 5'-AACTATACAACCTACTACCTCA.

22G-1: 5'-AAAGTGGTCAAGCACGGTTAAC

22G-2: 5'-AGTAAACCCAGCTTTCTTGTAC

22G-1 and 22G-2 probes were mixed before hybridization in Ambion hybridization buffer (ULTRAhyb-Oligo). Blots were exposed to phosphor-imager screens that were scanned on a BAS-2500 imager.

REFERENCES

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150, 88–99.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M.E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E. a (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science (80-.). 337, 574–578.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31, 67–78.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322, 1387–1392.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489, 447–451.

Cecere, G., Zheng, G.X.Y., Mansisidor, A.R., Klymko, K.E., and Grishok, A. (2012). Promoters recognized by forkhead proteins exist for individual 21U-RNAs. Mol. Cell 47, 734–745.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell 31, 79–90.

Duchaine, T.F., Wohlschlegel, J. a., Kennedy, S., Bei, Y., Conte, D., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G., et al. (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124, 343–354.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94–108.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151, 1488–1500.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D. V, Blaser, H., Raz, E., Moens, C.B., et al. (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129, 69–82.

Huang, X.A., Yin, H., Sweeney, S., Raha, D., Snyder, M., and Lin, H. (2013). A major epigenetic programming mechanism guided by piRNAs. Dev. Cell 24, 502–516.

Kamminga, L.M. (2011). RNAi genes pave their own way. Gene 529–533.

Kamminga, L.M., van Wolfswinkel, J.C., Luteijn, M.J., Kaaij, L.J.T., Bagijn, M.P., Sapetschnig, A., Miska, E.A., Berezikov, E., and Ketting, R.F. (2012). Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans. PLoS Genet. 8, e1002702.

Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3’ end formation of PIWI-interacting RNAs in vitro. Mol. Cell 43, 1015–1022.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150, 78–87.

Luteijn, M.J., and Ketting, R.F. (2013). PIWI-interacting RNAs: from generation to transgenerational epigenetics. Nat. Rev. Genet. 14, 523–534.

Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO J. 31, 3422–3430.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell 136, 656–668.

Olivieri, D., Sykora, M.M., Sachidanandam, R., Mechtler, K., and Brennecke, J. (2010). An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. EMBO J. 29, 3301–3317.

Patel, D.J., Ma, J.-B., Yuan, Y.-R., Ye, K., Pei, Y., Kuryavyi, V., Malinina, L., Meister, G., and Tuschl, T. (2006). Structural biology of RNA silencing and its functional implications. Cold Spring Harb. Symp. Quant. Biol. 71, 81–93.

Pavelec, D.M., Lachowiec, J., Duchaine, T.F., Smith, H.E., and Kennedy, S. (2009). Requirement for the ERI/DICER complex in endogenous RNA interference and sperm development in Caenorhabditis elegans. Genetics 183, 1283–1295.

Rozhkov, N. V, Hammell, M., and Hannon, G.J. (2013). Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev. 27, 400–412.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127, 1193–1207.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150, 65–77.

Sienski, G., Dönertas, D., and Brennecke, J. (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151, 964–980.

Smibert, P., Yang, J.-S., Azzam, G., Liu, J.-L., and Lai, E.C. (2013). Homeostatic control of Argonaute stability by microRNA availability. Nat. Struct. Mol. Biol. 20, 789–795.

Spencer, W.C., Zeller, G., Watson, J.D., Henz, S.R., Watkins, K.L., Mcwhirter, R.D., Petersen, S., Sreedharan, V.T., Widmer, C., Jo, J., et al. (2011). A spatial and temporal map of C . elegans gene expression. 325–341.

Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A. a, and Tóth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. 27, 390–399.

Waaijers, S., Portegijs, V., Kerver, J., Lemmens, B.B.L.G., Tijsterman, M., van den Heuvel, S., and Boxem, M. (2013). CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. Genetics 195, 1187–1191.

Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. Curr. Biol. 18, 861–867.

SUPLEMENTAL INFORMATION:

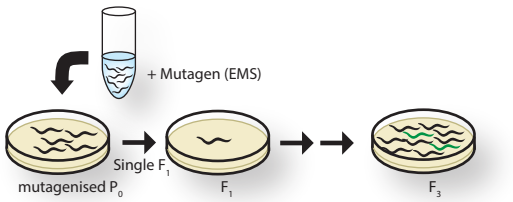
Supplemental table 1. Sequencing statistics of analyzed libraries.

Library	Total reads	Mapped	rRNA	miRNA	21U RNA	21U/miRNA
wild-type sensor TAP	28511907	24906523	9226784	7026459	480372	0.07
WT sensor OX	25404121	20201525	456461	6356515	6870671	1.1
xf14 sensor TAP	30621020	24873273	6014967	8877665	45552	0.005
xf14 sensor OX	19170339	11683426	1072921	6617227	701281	0.1
pid-1(xf14) CIP-TAP	11570860	2057187	707648	9041	4443	0.5
pid-1(x14) TAP	17995314	13653587	4482828	2615263	15272	0.006
pid-1(xf35) TAP	22421443	15883626	8896204	1290701	8801	0.007
pid-1(xf36) TAP	23142408	16766007	9266675	1513426	10664	0.007
wild-type CIP-TAP	9735643	4476441	1839775	24363	26703	1.1
wild-type TAP	22024861	17817074	4191309	2659164	393590	0.2
prg-1(pk2298) CIP-TAP	9555880	5945790	1359036	14941	19864	1.3
prg-1(pk2298) TAP	22882756	18311813	4869714	5269031	9066	0.002

Supplemental table 2. List of non-synonimous mutations found *pid-1(xf14)* worms.

gene	amino-acid substitution	gene	amino-acid substitution
F22G12.3	A308T	ZK593.5	G731E
F47G6.2	I566N	C08F11.10	M1I
M01E11.3	E470K	Y37A1B.2	G92S
F36H2.5	V227I	Y105C5B.25	H174Y
ZK970.6	A157T	Y38F2AL.5	I115S
F37H8.3	E160K	F28E10.4	E186K
E01G4.5	G57R	H06H21.10	M615I
K04B12.1	V1069A	M03D4.1	Y302F
R52.7	M46I	F38A5.13	P473S
C17F4.3	M159I	T22D1.2	T105S
F18A1.8	R61C	F45E4.4	A2310V
ZK1320.5	P359L	C08G9.2	G720E
K10G9.1	P446S	C18F3.4	M1I
D2045.2	G1025S	T11B7.4	V1102D
Y66D12A.9	H192Y	F56D5.3	E373K
Y56A3A.16	G133R	F56E10.1	L197F
C14B1.9	R89C	T01D3.1	Q14*
C14B1.9	R89C	AH10.1	G324E
R10E9.2	C137Y	C31A11.5	G319E
C26E6.8	L242F	Y6G8.14	N77K
R01H2.3	T503I	K02E2.7	M62I
K04G7.1	G428S	T03D8.1	A677T
K04G7.1	G428S	C02E7.9	P11S
F56C9.10	P171L	R07B5.7	G103E
ZK688.7	A58V	K08H10.7	L282F
C30A5.2	G193E	VZC374L.1	P51S
F22B7.13	A234T	C23H4.6	L615F
K04H4.2	G701R	K06A9.2	P482L
K08F4.2	V352M	H02F09.3	S1206L
K08F4.8	G43R	C18A11.5	H90Y
R09H10.4	S845L	F41D9.5	G694E

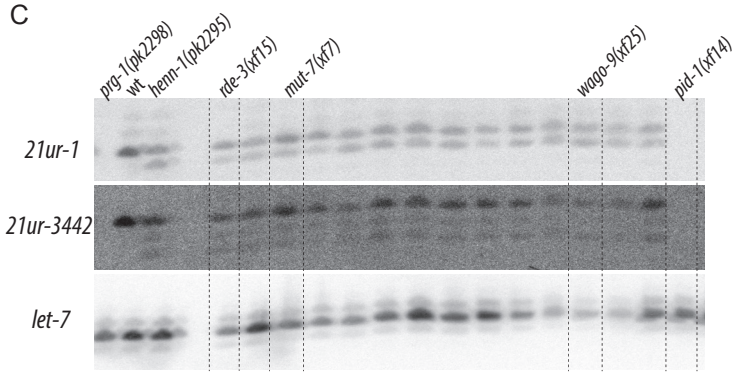
A



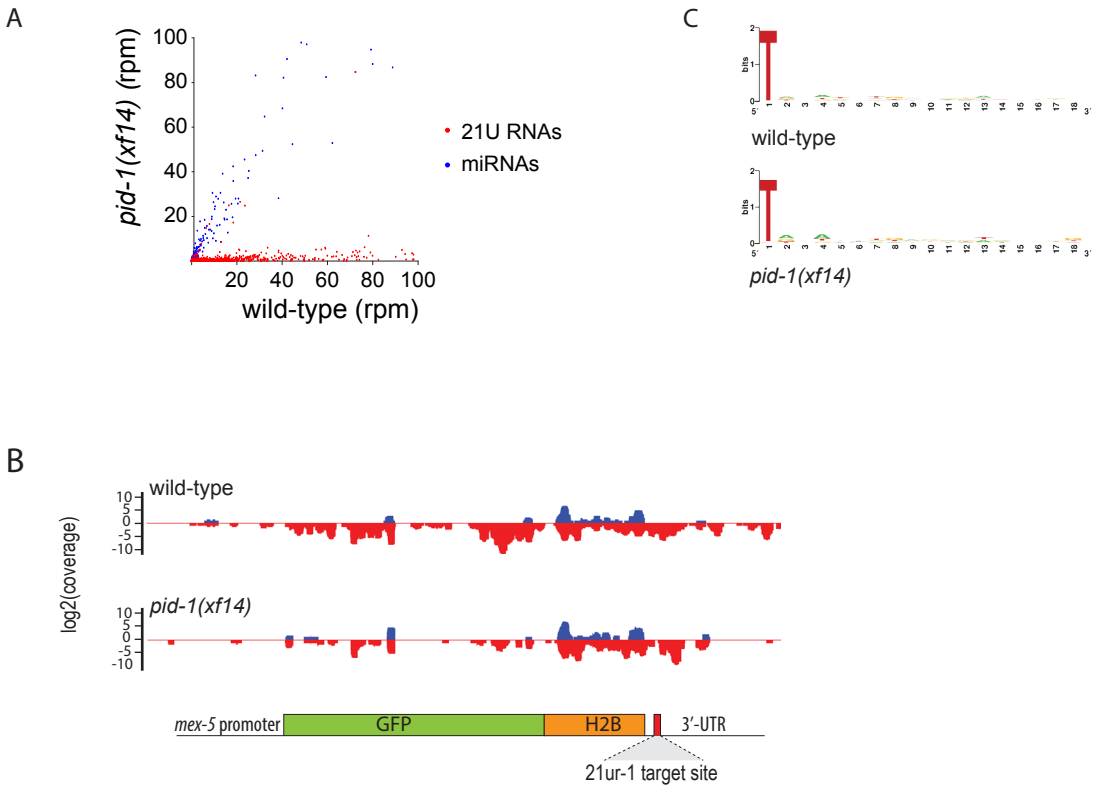
B

Allele	Gene	Mutation
xf7	*mut-7*	Q131STOP
xf15	*rde-3*	Q146STOP
xf25	*wago-9/hrde-1*	A1032T

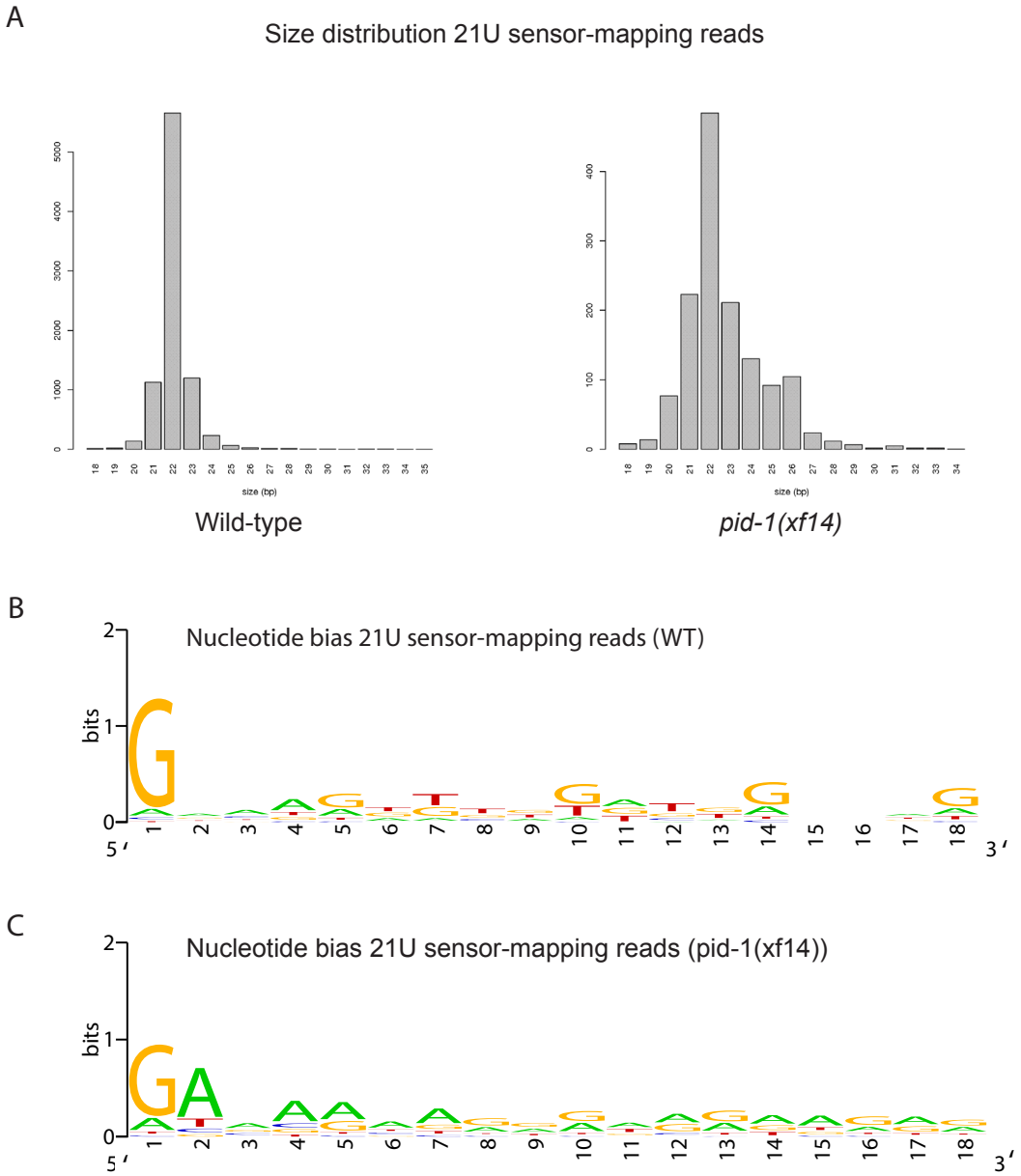
C



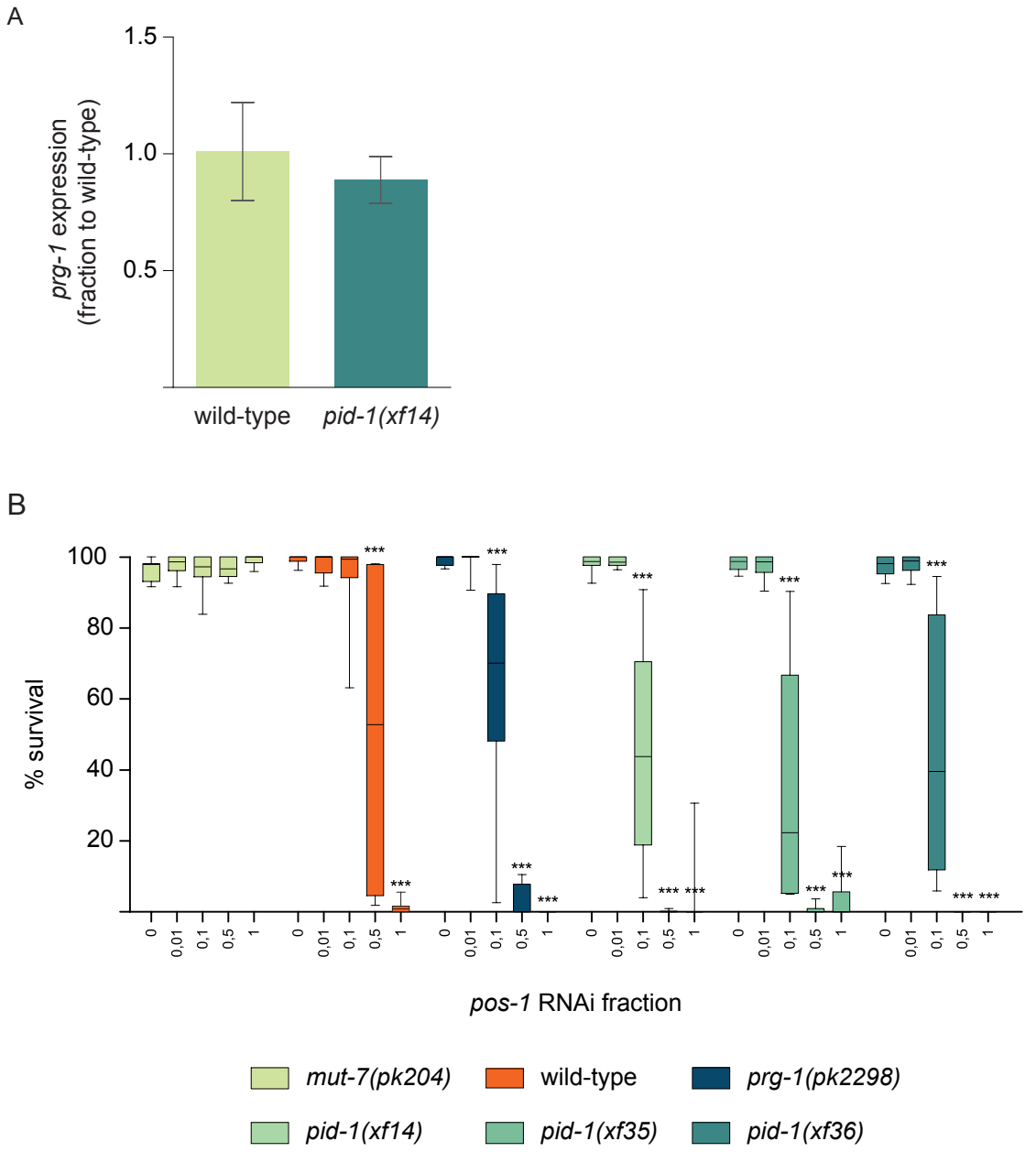
Supplemental figure 1. Isolation of 21U pathway mutants. A) Schematic of screen. Animals carrying a mutant the *henn-1(pk2295)* and a sensor that is silenced by *21ur-1* is mutagenized using EMS. F1 progeny was cloned out and was screened in the F3 for reactivation of the sensor. Isolated mutants were named Pid, for Piwi-Induced silencing defective. **A)** Identified alleles of factors known to act in the 21U pathway. **C)** Northern blots probed for two 21U RNA species (*21ur-1* and *21ur-3442*). Total RNA of isolated mutants was loaded. Non-marked lanes represent mutants not further discussed here. All mutants still harbor the *henn-1(pk2295)* mutation, causing the doublet signal of the 21U RNAs. The miRNA *let-7* is used as loading control.



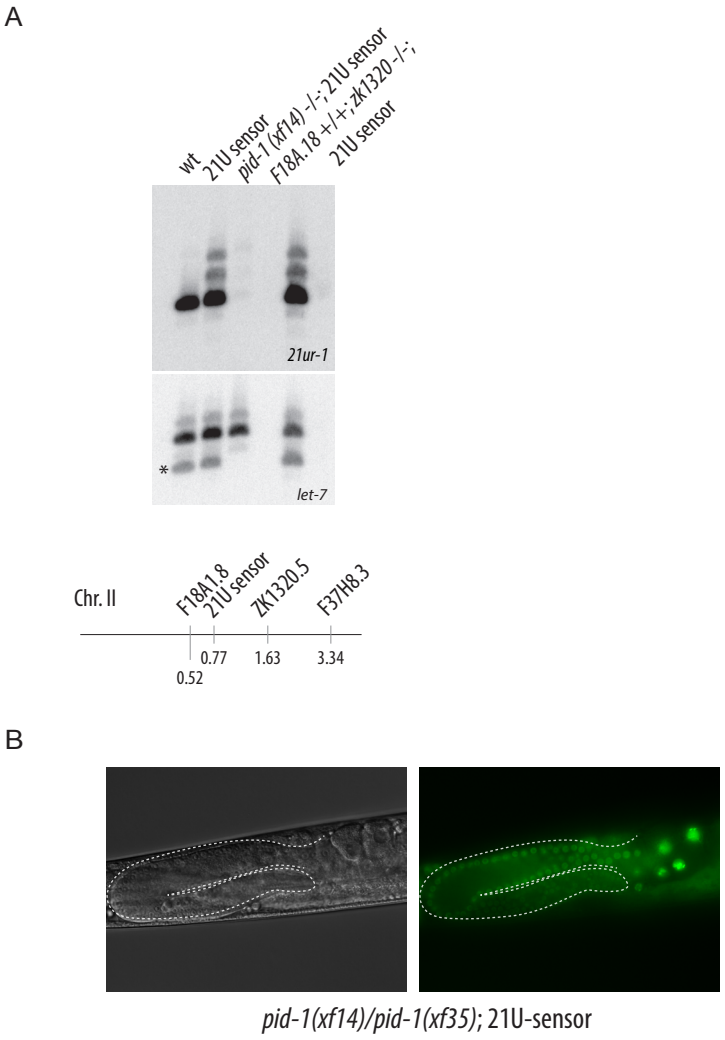
Suplemental figure 2. A) scatter plots of comparing *pid-1(xf14)* and wild-type small RNA-libraries. A) Scatter plot comparing miRNA and 21U RNA cloning frequencies between wild-type and *pid-1(xf14)* mutant animals. **B)** Normalized coverage (log2 transformed) of the 21U sensor by 22G RNAs. Red: antisense, Blue: sense. **C)** Sequence logos of 21U RNA-reads from wild-type and *pid-1(xf14)* mutant cDNA libraries.



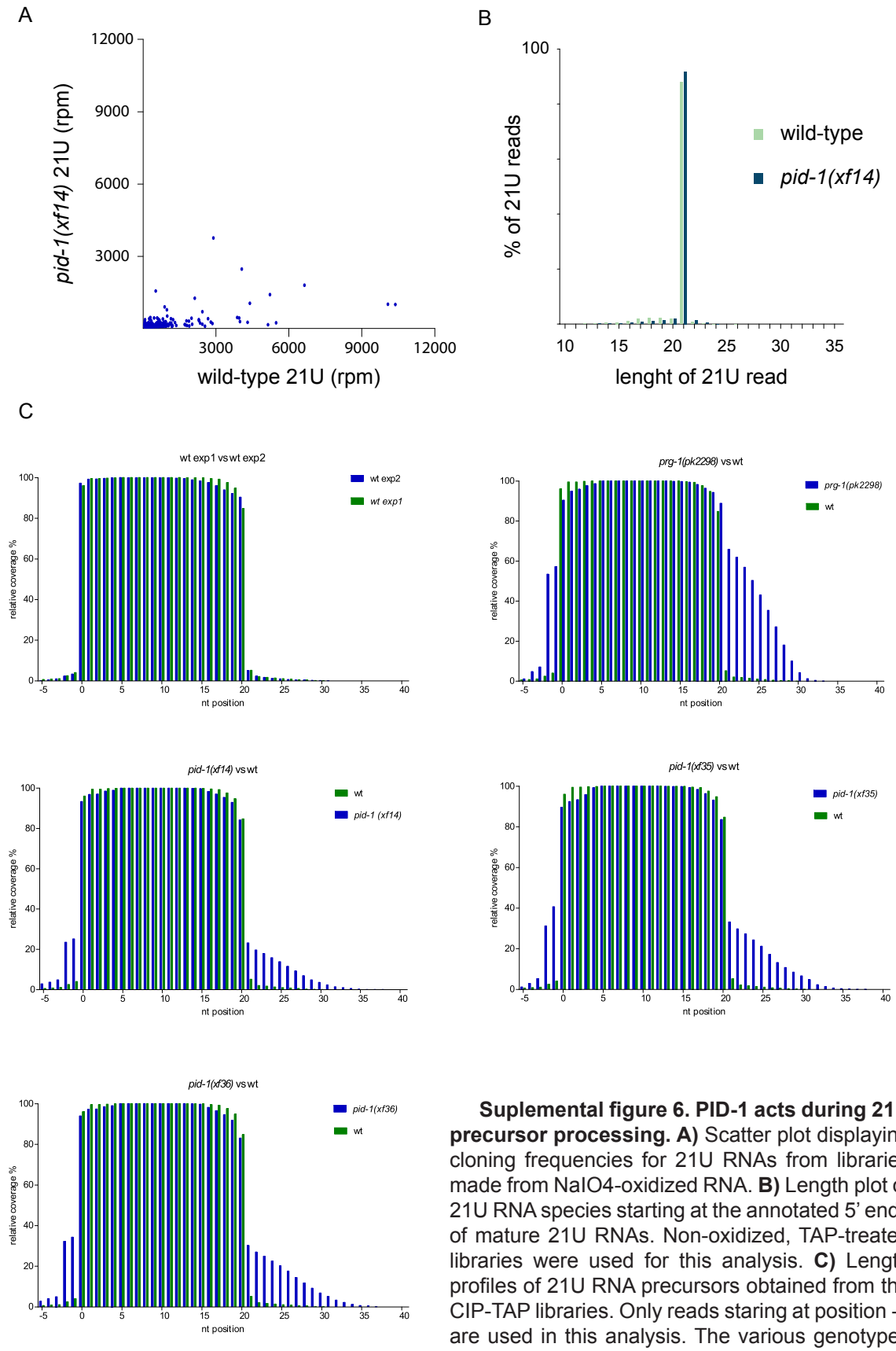
Supplemental Figure 3. Characteristics of sensor-derived small RNAs. A) Size distribution of transgene-mapping small RNAs from the two indicated genotypes. Both strains are homozygous for the 21U-sensor transgene. **B)** Sequence logo of transgene-mapping RNAs from wild-type animals. **C)** Sequence logo of transgene-mapping RNAs from *pid-1(xf14)* animals.



Suplemental figure 4. Prg-1 levels and germline RNA resistance in *pid-1* mutants A) RT-qPCR analysis of prg-1 transcript levels in wild-type and pid-1(xf14) mutant animals. B) RNAi sensitivity to pos-1 expressing E. coli was tested by diluting (10-fold and 100-fold) the RNAi-inducing food with non-transgenic OP50 bacteria. Survival reflects resistance against RNAi. Asterisks indicate statistically significant changes ($p<0.0001$, two-way Annova test, Bonferroni corrected)



Supplemental figure 5: Identification of *pid-1*. A) RNA from different recombinants isolating the mutations found in the indicated genes, was probed for the presence or absence of 21ur-1 RNA using Northern blotting. Indicated strains are wild-type for *henn-1*. The asterisk labels remaining 21U signal after stripping and re-probing for *let-7*. **B)** Complementation test between the *xf14* and *xf35* alleles of *pid-1*. GFP expression was scored in animals that had been trans-heterozygous for two generations.



Suplemental figure 6. PID-1 acts during 21U precursor processing. A) Scatter plot displaying cloning frequencies for 21U RNAs from libraries made from NaIO4-oxidized RNA. **B)** Length plot of 21U RNA species starting at the annotated 5' ends of mature 21U RNAs. Non-oxidized, TAP-treated libraries were used for this analysis. **C)** Length profiles of 21U RNA precursors obtained from the CIP-TAP libraries. Only reads staring at position -2 are used in this analysis. The various genotypes used are indicated in the figure.

Chapter 3.1

PID-1 interacts with IFE-3 and TOFU factors

Bruno F.M. de Albuquerque, Falk Butter and René F. Ketting

Following up on *pid-2*

In chapter 3 we described the identification of PID-1, a novel factor involved in biogenesis of piRNAs (de Albuquerque et al., 2014). However, the precise function of PID-1 in the piRNA biogenesis is still unclear. To gain further insights about the role of PID-1 in the biogenesis of piRNAs we performed comparative proteomics to identify factors that interact with PID-1.

PID-1 interacts with IFE-3 and TOFU factors

We immuno-precipitated PID-1 from wild-type, *pid-1(xf14)* and *pid-1(xf35)* mutant worms and analyzed the immune-precipitates by label free mass-spectroscopy. We then determined the enrichment of each protein for wild-type animals in respect to pid-1 mutants (figure 1, table 1). Both analysis show similar results and we found that pid-1 interacts with TOFU-6, Y23H5A.3, F35G12.11 and IFE-3.

TOFU-6 (Twenty-One-u Fouled Up -6), also known as MEL-47 (Maternal Effect Lethal-47), contains a tudor domain and a RNA binding domain (RRM) and is required for piRNA biogenesis (Goh et al., 2014). Knockdown of TOFU-6 results in decreased levels of mature piRNAs while maintaining normal levels of piRNA precursors. This is consistent with what we observed for *pid-1* mutants, where we saw a relative accumulation of piRNA precursors (de Albuquerque et al., 2014). However, in *pid-1* mutants, we also see a general decrease in precursor levels which that *pid-1* may also be involved in a step upstream of TOFU-6.

Y23H5A.3 contains a DNA binding domain and knockdown of Y23H5A.3 using RNAi, results in a 2 fold depletion of mature piRNAs (Goh et al., 2014). Interestingly Y23G5A.3 was shown to interact with TOFU-6 in a Y3H screen for TOFU-6 interactors (Minasaki and Streit, 2007).

IFE-3 is one of the five C. elegans homologs of the mRNA cap-binding protein eIF4E, and in-vitro studies shown that it binds monomethylated guanosine cap structures (m^7-GTP) but does not bind to trimethylated guanosine cap structures ($m_3^{2,2,7}$-GTP) (Ruszczyńska-Bartnik et al., 2011). In *C. elegans* around 70% of the protein coding genes are trans-spliced to one of two 22 nucleotides splice leaders (Riddle et al., 1997) that contain a $m_3^{2,2,7}$-GTP 5'-CAP, whereas the remaining 30% of the protein coding genes are not trans-spliced and have a m^7-GTP 5'-CAP. In

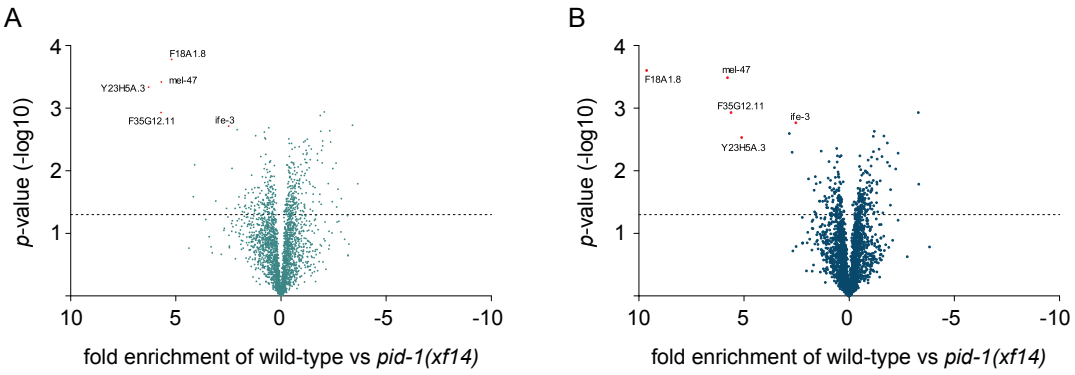


Figure. PID-1 IP-MS. Volcano plot depicting peptide enrichment vs *p*-value (Welch t-test) of between PID-1 immuno-precipitates of wild-type and: **A)** *pid-1(xf14)* mutants and **B)** *pid-1(xf35)* mutants. Dotted lines represent a p-value of 0.05. Red dots represent factors that are commonly enriched in both analyses.

Table 1. Proteins enriched in PID-1 immuno- complexes

wild-type vs pid-1(xf14)

MS/MS Count	*p*-value -log10	Welch Difference	Gene name
266	3,779250287	-5,193398952	F18A1.8
137	3,418015959	-5,684963703	mel-47
175	3,336865093	-6,288874149	Y23H5A.3
62	2,930076975	-5,699674129	F35G12.11
124	2,713756042	-2,490641117	ife-3
22	2,65536691	-2,083729744	rmd-1
33	2,094245608	-4,103991985	lbp-9
41	2,038185314	-2,326582909	prmt-1
56	1,584504914	-4,162981987	nab-1
29	1,51589477	-3,094854832	pud-1,2
7	1,459586091	-2,019426823	lbp-5
39	1,399053152	-2,58891964	fah-1
6	1,398256762	-2,179341793	gln-6

wild-type vs pid-1(xf35)

MS/MS Count	*p*-value -log10	Welch Difference	Gene name
266	3,60239117	-9,631870747	F18A1.8
137	3,487650969	-5,787314415	mel-47
62	2,93063305	-5,614844799	F35G12.11
175	2,530185367	-5,114140987	Y23H5A.3
6	2,59466032	-2,844799519	gln-6
10	2,29483204	-2,710696697	dhs-27
124	2,766372223	-2,540003777	ife-3

C. elegans, piRNA precursors do not undergo trans-splicing (Gu et al., 2012). The specificity of EFI-3 for m7GTP and the fact that it interacts with PID-1 suggest that EFI-3 is indeed bound to the precursor 5'-CAP.

F35G12.11 is a novel protein that contains an Enhancer of Rudimentary domain, which is a highly conserved domain in animals and plants, but so far with unknown function.

Interestingly, three of mentioned factors, TOFU-6, Y23H5A.3 and IFE-3, have a MEL phenotype, which we don't see in either *pid-1* or *prg-1* mutants. This suggests that these factors act together in mRNA processing, and their role is not confined to piRNA biogenesis.

Final remarks

So far we were unable to see any interaction between PID-1 and piRNA precursors, either by immune-precipitation or in-vitro studies (not shown). The fact PID-1 interacts with at least three factors that have a RNA binding domain, one of them with affinity to a 5'-CAP structure specific to piRNA precursors, suggest that PID-1 does not bind to the precursors. The different levels of piRNA precursors between *pid-1* mutants and worms treated with RNAi against TOFU-6, suggests that *pid-1* also plays a role in precursor stability upstream TOFU-6 and Y23H5A.3, which is in line with our initial prediction that PID-1 may be involved in transport of piRNA precursors to the cytoplasm (de Albuquerque et al., 2014).

REFERENCES

De Albuquerque, B.F.M., Luteijn, M.J., Cordeiro Rodrigues, R.J., van Bergeijk, P., Waaijers, S., Kaaij, L.J.T., Klein, H., Boxem, M., and Ketting, R.F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28, 683–688.

Goh, W.S., Seah, J.W.E., Harrison, E.J., Chen, C., Hammell, C.M., and Hannon, G.J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. Genes Dev. 28, 797–807.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151, 1488–1500.

Minasaki, R., and Streit, A. (2007). MEL-47, a novel protein required for early cell divisions in the nematode Caenorhabditis elegans. Mol. Genet. Genomics 277, 315–328.

Riddle, D.L., Blumentha, T., Meyer, B.J., and Priess, J.R. (1997). C. elegans II. 2nd edition. Cold

Spring Harb. Cold Spring Harb. Lab. Press.

Ruszczyńska-Bartnik, K., Maclejczyk, M., and Stolarski, R. (2011). Dynamical insight into Caenorhabditis elegans eIF4E recognition specificity for mono-and trimethylated structures of mRNA 5′ cap. J. Mol. Model. 17, 727–737.

Chapter 3

Maternal piRNAs are essential for germline development following de-novo establishment of endo-siRNAs in *Caenorhabditis elegans*.

Bruno F.M. de Albuquerque, Maria Placentino and René F. Ketting

SUMMARY

The Piwi-piRNA pathway represents a germline specific transposon-defense system. *C. elegans* piwi, *prg-1*, is a non-essential gene and triggers a secondary RNAi response that depends on so-called mutator genes, endo-siRNAs (22G-RNAs) and at least one 22G-RNA-binding Argonaute protein, HRDE-1. Interestingly, through a poorly understood mechanism, silencing of PRG-1 targets can become PRG-1 independent. This state, also known as RNAe, is heritable and depends on mutator genes and HRDE-1. We studied how the transgenerational memory of RNAe and the piRNA pathway interact. We find that after elimination of mutator-dependent 22G-RNAs, PRG-1 is required for their re-establishment. Strikingly, attempts to re-establish this pool in absence of PRG-1 result in severe germline proliferation defects, accompanied by off-target mutator-activity and a disturbed balance between gene-activating and repressing 22G-RNA pathways. Our results demonstrate that maternal piRNAs in C. elegans are required to initiate transposon-targeting 22G-RNA pathways and to prevent ectopic activity of HRDE-1-mediated silencing.

INTRODUCTION

The Piwi-piRNA pathway is an RNAi-related mechanism that is essential for germ cell development in most organisms (Ghildiyal and Zamore, 2009; Ketting, 2011; Malone and Hannon, 2009). Loss of this pathway results in strong up-regulation of transposon activity, apoptosis and blocks at various stages of meiosis. In addition, in zebrafish loss of one of the Piwi proteins results in a lack of germ cell proliferation (Houwing et al., 2008). Hence, the Piwi-piRNA pathway is considered as an essential feature of germ cells in general, and well-known germ cell markers, such as Vasa and Tdrd1, are in fact integral parts of the molecular mechanisms of the Piwi pathway.

In contrast, the *C. elegans* Piwi pathway has been shown to be not acutely required in germ cells. Loss of the Piwi protein PRG-1 results only in relatively minor problems, including defects in spermatogenesis and a reduced brood-size (Batista et al., 2008; Cox et al., 1998; Das et al., 2008; Wang and Reinke, 2008). Importantly, the impact of PRG-1 on transposon silencing appears to be only marginal, as one transposon was shown to be activated upon loss of PRG-1 (Das et al., 2008). Intriguingly, while *prg-1* mutant animals are fertile, they do have a so-called mortal germline (Mrt) phenotype (Simon et al., 2014), meaning that the germline tends to be lost over the course of generations. This does not seem to relate to genomic damage, as is often observed in Piwi mutants in other species, but rather to an accumulation of epigenetic defects (Simon et al., 2014).

Mechanistically, PRG-1 triggers the production of endo-siRNAs (22G RNAs). This occurs in a process that depends on an RNA dependent RNA polymerase, even though this has not been directly demonstrated, and so-called mutator proteins (Bagijn et al., 2012; Lee et al., 2012). Mutator proteins act in a complex that contains the RdRP enzyme RRF-1 and are required for the accumulation of many 22G-RNA species in many different RNAi-related pathways (Zhang et al., 2011). These proteins are physically concentrated at discrete sites, named mutator-foci (Phillips et al., 2012) that flank bigger cytoplasmic, peri-nuclear aggregates known as P-granules. Animals lacking one of these mutator genes display general defects in RNAi and activation of many different transposable elements (Ketting et al., 1999; Tabara et al., 1999). Different Argonaute proteins likely act as recipients for the 22G RNA output of mutators. These include WAGO-1 (Gu et al., 2009), PPW-1 (Simon et al., 2014) and HRDE-1 (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al.,

2012; Shirayama et al., 2012). Interestingly, while *hrde-1* mutants also display a Mrt phenotype (Buckley et al., 2012), mutator mutants by themselves do not (Simon et al., 2014). However, mutator genes do affect germline mortality in a certain genetic background: in *prg-1;daf-2* double mutants the Mrt phenotype of *prg-1* is suppressed, and mutator genes are required for this suppression (Simon et al., 2014).

As mentioned, PRG-1 can silence target genes through the involvement of mutator genes. Interestingly, such PRG-1-initiated silencing can become PRG-1 independent (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). This state, referred to as RNAe, can be faithfully inherited across many generations and fully depends on mutators, the nuclear 22G-RNA-binding Argonaute protein HRDE-1 and chromatin factors (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). This indicates that mutator-mediated silencing, at least on a transgenic target, needs PRG-1 to start, but can become self-sustainable. This can be seen as a form of trans-generational silencing memory. In this light it is interesting to note that transposon activation is much stronger in mutator mutants than in *prg-1* mutants, as one of the most active transposons in *C. elegans*, *Tc1*, is still fully silent in *prg-1* mutants (Das et al., 2008). Possibly, transposon silencing, at least for some transposons, depends for a large extent on the PRG-1-independent memory, and does not need to be initiated each generation.

In parallel to a memory that transmits silencing, *C. elegans* gametes also transmit information on genes that are active (Conine et al., 2013; Seth et al., 2013; Wedeles et al., 2013). This requires the Argonaute proteins ALG-3, ALG-4 and CSR-1. In fact, CSR-1 can re-activate genes that have been silenced through PRG-1 and mutator activity (Seth et al., 2013). The molecular mechanisms behind this activation of gene expression are currently not clear. These may involve chromatin-related effects (Claycomb et al., 2009; Wedeles et al., 2013), but could also relate to 22G-RNA turnover, since we previously described an enzyme named CDE-1 that is required for CSR-1-bound 22G RNA turnover through non-templated uridylation of specifically these 22G-RNAs (van Wolfswinkel et al., 2009).

We set out to directly test the idea that transposons are kept silenced through RNAe-related memory and that indeed PRG-1 is required for the initiation of transposon silencing only. To do this we first erased the mutator mediated silencing memory from *prg-1* mutant animals, and then reactivated this memory system. Indeed, we find that transposon-targeting 22G RNAs require PRG-1 to re-establish, and we demonstrate that PRG-1 is required to silence *Tc1* in an *hrde-1* mutant

background. In addition, however, these crosses revealed an acute requirement for PRG-1 in germ cell proliferation in early larva. We propose that this defect is related to mis-targeted mutator complexes that start to act, through HRDE-1, on transcripts that are normally protected from silencing through CSR-1.

RESULTS

Mutator-induced sterility in *prg-1* mutants

To erase RNAe-related memory from *prg-1* mutants, we created two strains (*prg-1; mut-7* and *prg-1;mut-16*) that lack both piRNAs (in *C. elegans* also named 21U RNAs) and RNAe. These lines exhibit strong transposon mobilization (not shown), and are fertile. Upon crossing these two lines with each other, offspring remains *prg-1* defective, while mutator activity will be re-established, allowing one to address whether *prg-1* is required to initiate transposon silencing (Figure 1A). Unexpectedly, however, the offspring of these crosses are completely sterile (Figure 1B, 1C, Figure S1). We first checked the generality of this phenotype by building a *prg-1* mutant strain that is also double mutant for *mut-14* and *smut-1*, two redundant RNA helicases (Phillips et al., 2014) that we now show are also required for RNAe maintenance (Figure 1D). This strain also yields sterile offspring when crossed with another mutator mutant lacking a functional *prg-1* pathway (Figure 1C). In addition, *prg-1* mutant animals in which *mut-7* and *mut-16* are complemented while *mut-14* and *smut-1* are kept homozygous mutant are fully fertile (Figure 1C). Finally, in these crosses *prg-1* mutations can be substituted by mutations in *pid-1*, a gene required for 21U-RNA biogenesis (de Albuquerque et al., 2014) (not shown). Collectively, these data show that the *prg-1* pathway is required to prevent sterility that is induced through de-novo establishment of mutator activity.

Mutator-induce sterility affects germ cell proliferation

We next probed the nature of the germ cell defect. Overall, the germ cell count is strongly reduced. Interestingly, both gonad arms tend to contain similar numbers of germ cells (Figure 2A), suggesting that the specific fate of germ cells in each individual is set rather early in development. The germ cells that are still present in these animals still express the germ cell marker PGL-1 (Kawasaki et al., 1998)

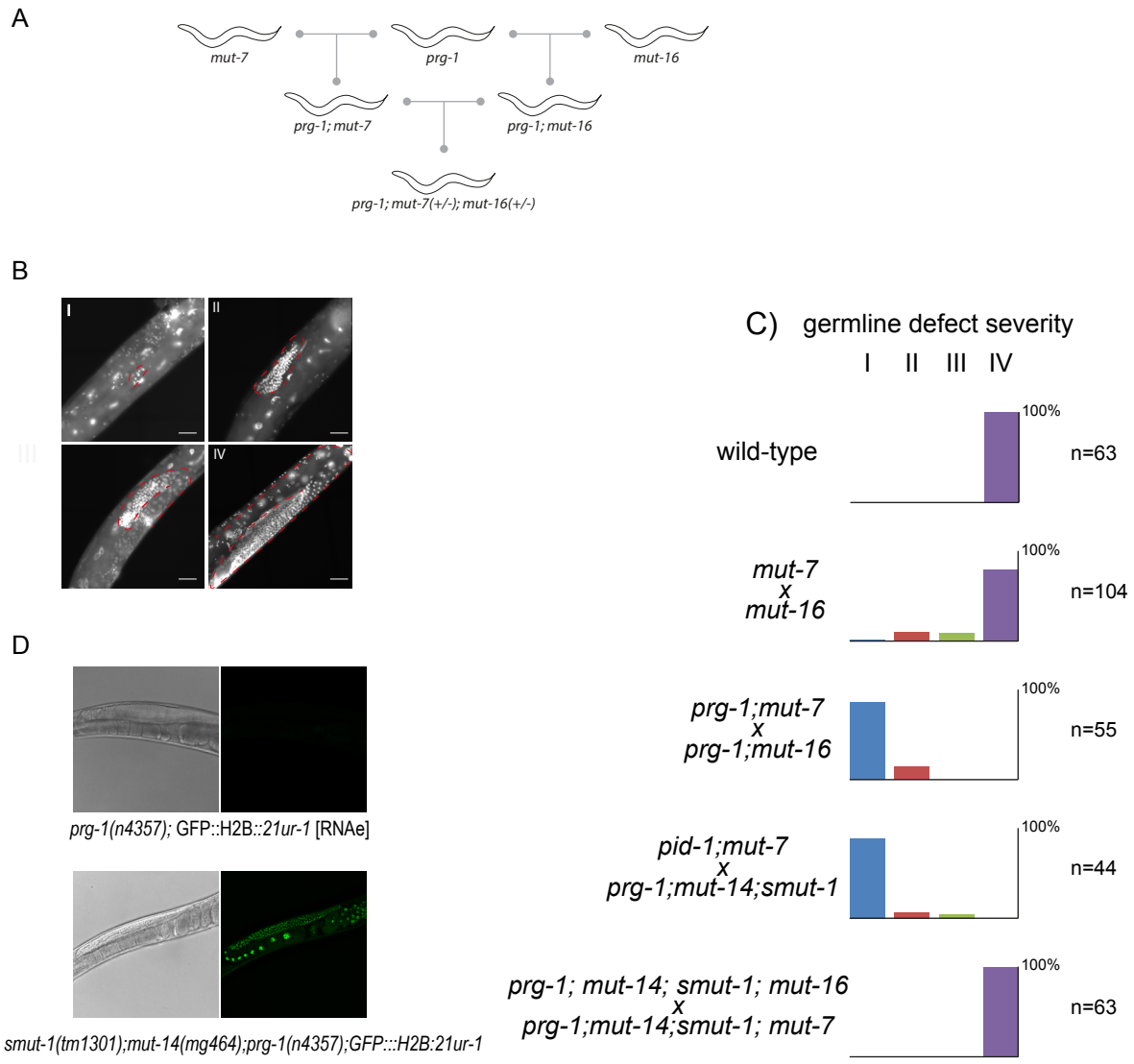


Figure 1. Mutator induced sterility. A) Schematic of the crosses performed to erase RNAe memory from *prg-1* mutant animals. Animals mutant for *prg-1* and distinct mutations for mutator genes lack both piRNA and RNAe activity. When crossed, mutations in the mutator genes complement each other re-establishing mutator activity in the absence of RNAe memory. **B)** DAPI staining of worms whose mutator activity was restored in the absence piRNA and RNAe memory. Gonads are outlined with a dashed line and scale bar represents 40 µm. The germline of individual animals was classified into one of four categories. Type I - none or very few germline cells; Type II - some germ cells but no apparent differentiation; Type III - some germ cells with apparent differentiation; Type IV – wild-type germline. **C)** Quantitation of the observed germline phenotypes in wild-type animals, animals were the mutator pathway was activated in the presence of a functional piRNA pathway (*mut-7* x *mut-16*) and animals were the mutator pathway was activated in a background with a disrupted piwi-pathway (*prg-1* or *pid-1* mutations). **D.** Confocal image of an epigenetically silenced 21U sensor that was reactivated in *smut-1;mut14* mutants. *smut-1* and *mut-14* mutant alleles were crossed into animals deficient for *prg-1* that carried an RNAe silenced 21U sensor. The resulting triple mutant is able to reactivate the sensor, whose silencing had previously become PRG-1 independent.

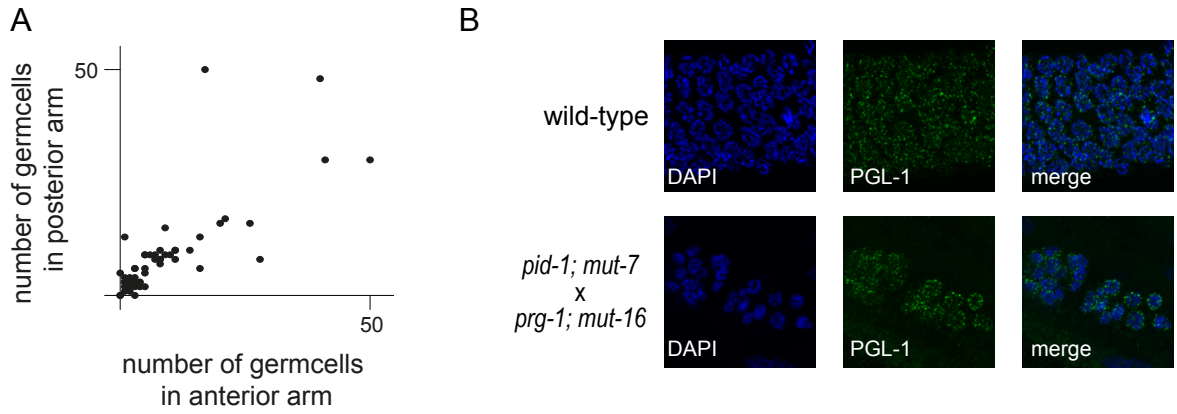


Figure 2. Characterization of Mutator-Induced-Sterility. A) The number of germ cells in both the posterior and the anterior arm were counted in worms displaying Mutator Induced Sterility (n=57). Mutator Induced Sterility was obtained by crossing *prg-1;mut-16*♀ with *pid-1;mut-7*♂. Both gonad arms tend to respond similarly. **B)** Confocal image of PGL-1 localization in animals displaying Mutator-Induced-Sterility. *prg-1;mut-16*♀ were crossed with *pid-1;mut-7*♂ and adult F1s were stained for PGL-1 (green) and DAPI (blue).

(Figure 2B). Likewise, we fail to detect expression of a somatic, neuronal gene unc-119 in the remaining germ cells (not shown), indicating that these germ cells are not subject to gross cell fate changes. Rather, the sterility is likely due to under-proliferation of germ cells during early larval development.

Maternal PRG-1 can rescue Mutator-induced sterility

We previously showed that *prg-1*-mediated silencing has a strong maternal component (de Albuquerque et al., 2014). We therefore asked whether maternal or paternal PRG-1 could rescue the sterility. For this, we first performed small RNA sequencing on three sets of L1 larva, in order to determine to what extent the 21U-RNAs are maternally derived. In L1 stage, the animals only have two germ cells, which can be considered the primordial germ cells of *C. elegans*, and these cells are not yet transcriptionally active. Interestingly, the levels of 21U-RNAs in offspring from a cross between wild-type males and *prg-1* mutant hermaphrodites are as low as those observed in straight prg-1 mutant L1 larva (Figure 3A). This shows that the vast majority of 21U-RNAs detected in wild-type L1 larva are of maternal origin. Consistent with this result, loss of maternal PRG-1 is sufficient to trigger sterility upon re-establishment of mutator activity (Figure 3B). We note that in contrast to *prg-1*, the mutator pathway needs to be mutated in both parents in order

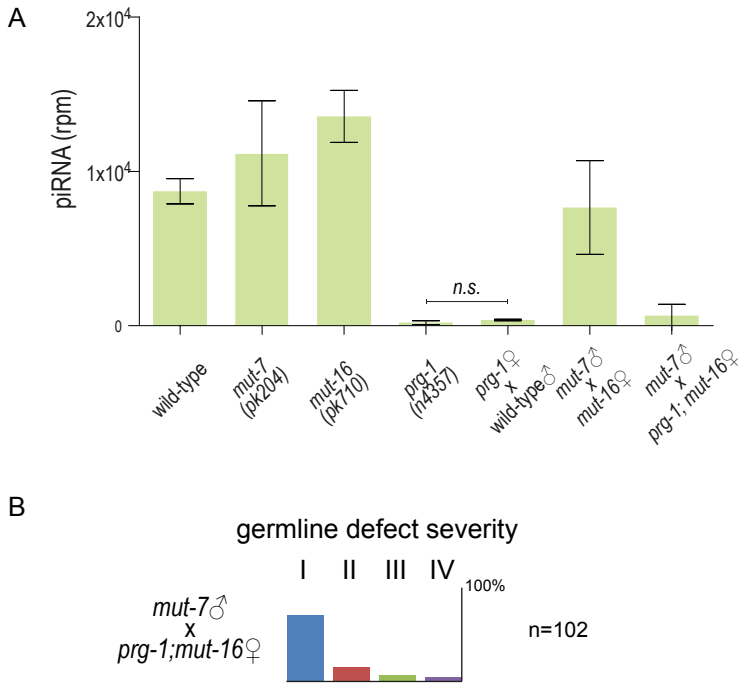


Figure 3. Maternal loading of 21U-RNA into offspring. A) Column chart showing piRNA levels determined by small RNA-seq of total RNA from L1 larva of the indicated genotype. piRNAs were considered as reads that were 20-21 nucleotide long, start with a U and were sense to annotated piRNA loci. Levels are indicated in “reads per million” of non-structural reads (rpm) and error bars represent standard deviation between at least two biological duplicates. Cross offspring L1 larvae derived from *prg-1(-/-)*♀ are depleted of piRNAs. **B)** Quantitation of the observed germline defects in the F1s of a cross between *prg-1(-/-);mut-16*♀ and *mut-7*♂. Activation of the mutator pathway in animals that only lack the maternal load of 21U RNAs results in sterility.

to trigger sterility, indicating that mutator genes transmit information both via sperm as well as via oocytes. Indeed, such paternal inheritance has been demonstrated (Stoeckius et al., 2014), even though RNAe-induced silencing has been shown to transmitted primarily through the female germline (Luteijn et al., 2012).

Initiation of Transposon-targeting 22G-RNAs requires PRG-1

We next analyzed small RNAs isolated from animals displaying mutator-induced sterility. Since the defect occurs early in development we focused our sequencing efforts on L1 and L2 larva. As controls, we used homozygous wild-type and corresponding mutant strains, as well as offspring of a cross between two mutator mutants that have wild-type PRG-1 activity (see Supplemental Information for an overview of the sequenced genotypes). For all samples two-three biological

replicates were processed, and progeny from crosses were hand-picked in order to make sure only cross-progeny was analyzed. Finally, since our samples were based on limited input material, we included four-nucleotide random barcodes at both 5' and 3' ends of the small RNA, in order to de-convolute PCR-amplification artifacts that can occur during library preparation.

We first looked at the abundance of transposon-targeting 22G-RNAs. These levels stay stable in wild-type animals, and in *prg-1* mutants, and, as expected, these 22G RNAs are largely missing in mutator mutants (Figure 4A). This shows that the majority of these 22G-RNAs are inherited in a PRG-1 independent manner, consistent with ongoing transposon silencing in *prg-1* mutants. When mutator activity is de-novo established in offspring from parents that both lack mutator activity, transposon-targeting 22G-RNAs start to build up in L2 stage. Interestingly, this only occurs when PRG-1 is present. In absence of PRG-1, no such build-up of transposon-targeting 22G-RNAs is observed (Figure 4A). This is good evidence that PRG-1 is required to initiate transposon silencing in *C. elegans*, while it is not required for the 22G-RNA mediated memory of it.

We also aimed to obtain more direct evidence for this idea. It has been shown previously that in *prg-1* mutants no activity of the *Tc1* transposon can be observed (Das et al., 2008). This was based on the lack of germline reversion of a *Tc1*-insertion in the muscle-expressed gene *unc-22*. We confirmed this, and also found that in *hrde-1* mutants, in which 22G-RNA memory should be compromised, *Tc1* is not de-silenced in the germline. In contrast, in mutator mutants, in which HRDE-1 cannot be fully active, *Tc1* is activated (Ketting et al., 1999; Tabara et al., 1999), suggesting that mutator activity loads other Argonaute proteins in addition to HRDE-1, and that these can maintain *Tc1*-silencing in *hrde-1* mutants. We reasoned, that the loading of such other Argonautes may be disturbed in *prg-1* mutants, and created a *prg-1;hrde-1* double mutant to check for *Tc1* activation, by probing the stability of a *unc-22::Tc1(st136)* insertion allele. We could detect *Tc1* excision events easily in this double mutant background (Figure S2), lending direct genetic support to the idea that in absence of trans-generational 22G-RNA memory, PRG-1 is required to silence *Tc1* transposition in the germline. It suggests that either PRG-1 itself directly silences *Tc1*, or that PRG-1 triggers the loading of additional Argonaute proteins besides HRDE-1, and that these affect *Tc1* activity. In the latter case, these additional secondary Argonaute proteins, unlike HRDE-1, require continued presence of PRG-1 to be effective.

Mutator-driven 22G-RNAs from CSR-1 target RNAs

In order to better understand why activation of mutator activity in a prg-1 background leads to sterility, we next checked the abundance of small RNAs that have been described to be enriched in particular Argonaute proteins, such as CSR-1, WAGO-1, ALG-3/4, PRG-1 and ERGO-1 (Figure 4A, Figure S3). While the different mutant background give variations from wild-type levels of 22G-RNAs in these pathways, none of these could easily explain the sterility, since loss or gain of 22G-RNAs of most of these pathways is known not to trigger sterility per-sé. The one exception is the CSR-1 pathway, and we next analyzed these 22G-RNAs in further detail.

CSR-1 has been proposed to sustain, rather than to silence gene expression (Claycomb et al., 2009; Seth et al., 2013; Wedeles et al., 2013), and has been shown to act primarily on germ cell-expressed genes. Interestingly, the RdRP enzyme EGO-1, which makes the CSR-1-bound 22G-RNAs, was originally identified as an enhancer of germ cell proliferation defects (Qiao et al., 1995), potentially linking the germline proliferation defect we observe, to CSR-1. It has been shown that CSR-1-bound 22G RNAs are mostly independent of mutator genes (Gu et al., 2009). In our data-sets we see this trend as well, although in three out of four L1-derived libraries, including the libraries made from animals that will become sterile, the level of CSR-1-pathway-associated 22G-RNAs is significantly reduced (Figure 4B), suggesting that mutator activity does contribute to part of the CSR-1 target-derived 22G-abundance. Still, this by itself cannot explain the sterility, since mutator mutants are not sterile.

22G-RNAs from individual genes are represented in multiple Argonautes

We next asked to what extent genes that are annotated as CSR-1 targets, are really exclusively targeted by CSR-1. In previously described 22G-RNA populations from immune-purified (IP) CSR-1 (Claycomb et al., 2009), one can detect significant enrichment of 22G-RNAs from genes that are not considered as typical CSR-1-target genes (Figure S4). Do these really represent CSR-1 bound 22G-RNAs or do they represent impurities in the IP-sample? To address that question we made use of the fact that non-templated uridylation of 22G RNAs, which is performed by the enzyme CDE-1, is a mark of the CSR-1 pathway (van Wolfswinkel et al., 2009). The

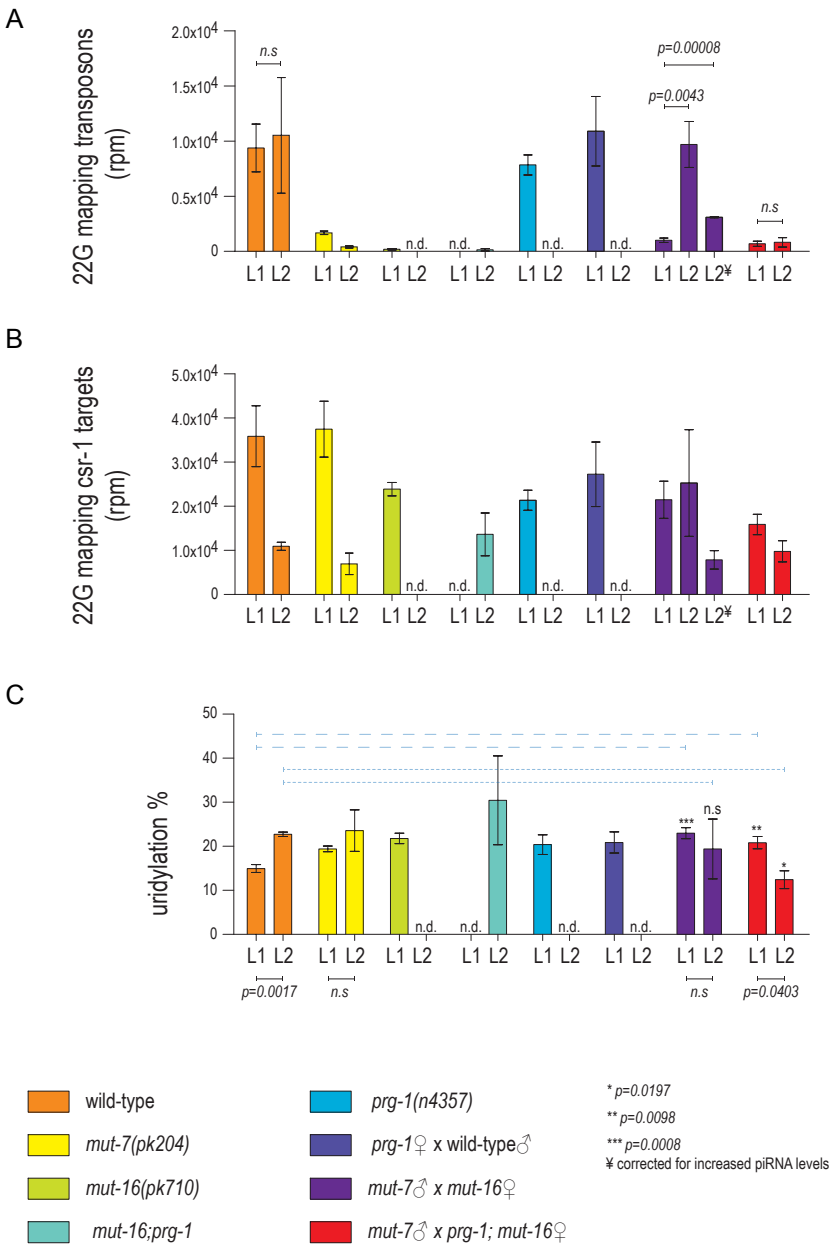
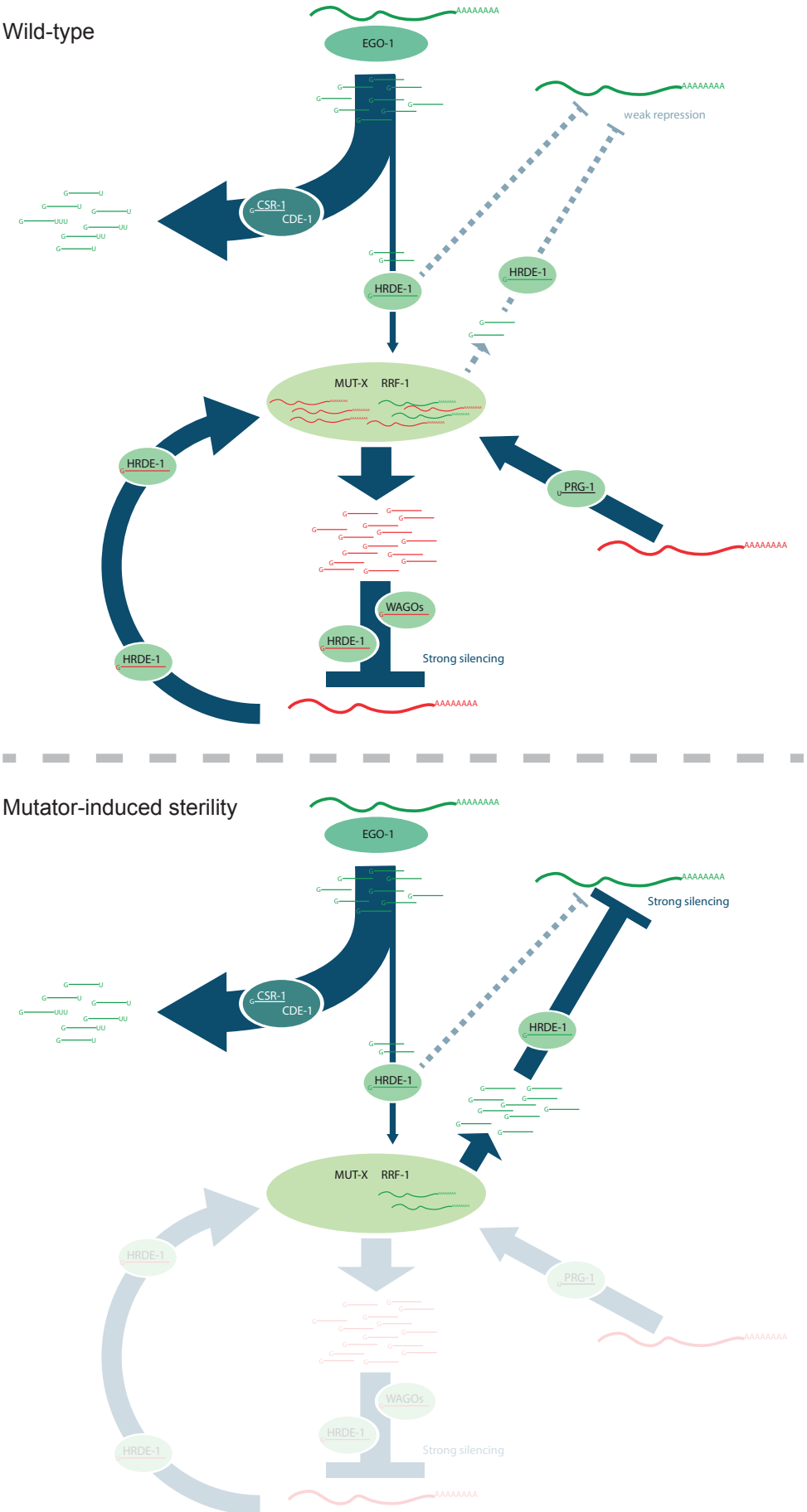


Figure 4. 22G-pathway analysis of Mutator-Induced-Sterility. A) Column chart showing levels of 22G small RNAs mapping anti-sense to transposons in L1 and L2 larvae of the indicated genotypes. 22G small RNAs were considered as reads that where 20-23 nucleotide long, start with a G and were anti-sense to annotated transposable elements. Levels are indicated in 'reads per million' of non-structural reads (rpm) and error bars represent the standard deviation between at least two biological duplicates. See Figure S3 for details on the correction marked with ¥. **B)** Column chart showing levels of 22G small RNAs mapping anti-sense to protein coding genes that are annotated as CSR-1 targets (TableS1). N.d. means 'not determined'. **C)** Fraction of non-templated uridylation of 22G-RNAs mapping protein coding genes targeted by CSR-1. P-values comparing L1 and L2 data of the same genotype are given below the x-axis. *p*-values for the comparison of different genotypes are given above the bars.

A


Wild-type
EGO-1
AAAAAAAA
CSR-1
CDE-1
weak repression
HRDE-1
MUT-X
RRF-1
PRG-1
WAGOs
Strong silencing
Mutator-induced sterility
EGO-1
AAAAAAAA
Strong silencing
CSR-1
CDE-1
HRDE-1
MUT-X
RRF-1
PRG-1
WAGOs
Strong silencing

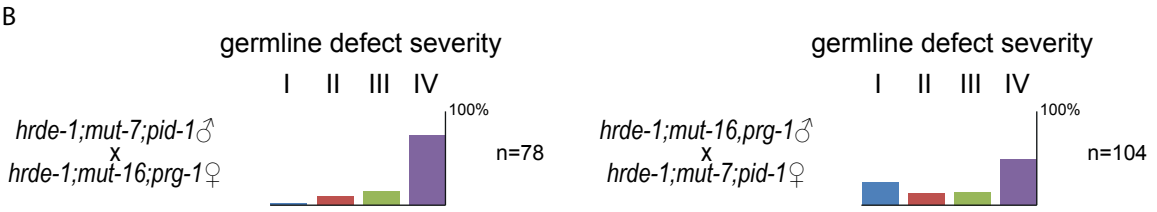


Figure 5. HRDE-1 is required for Mutator-Induced-Sterility. A) (oposite page) Model describing how mutator-induced sterility can develop. Under normal conditions, mutator activity is focussed on transcripts (Red) that are fed into mutator foci by PRG-1 and HRDE-1. This mutator activity stimulates RRF-1-mediated 22G RNA production and channels their loading into Argonaute proteins such as HRDE-1 and WAGO-1. These mediate effective silencing of their targets. Argonaute proteins such as HRDE-1 normally are also loaded to a minor extent by EGO-1. This is mostly inhibited through CSR-1/CDE-1-mediated 22G-RNA turnover. The small amounts of EGO-1 target (green) that do end up in mutator foci experience only mild mutator activity, since most of the mutator activity is consumed by PRG-1 and HRDE-1-mediated input. Hence, EGO-1 targets only experience little silencing and remain effectively expressed. In the scenario where mutator foci no longer receive input via PRG-1, nor from HRDE-1, mutator activity on EGO-1 targets will increase, leading to stronger silencing of EGO-1 targets. Since these targets typically are genes required for germ cell development, germ cell defects will develop. Loss of HRDE-1 rescues these defects because the silencing experienced by the EGO-1 targets diminishes. **B)** Quantification of the observed germline phenotypes in animals where the mutator pathway was reactivated in the absence of parental piRNAs, but also lacking HRDE-1.

22G-RNAs from the CSR-1 IP data that are derived from genes considered to be main targets of other Argonautes also show this increased uridylation rate (Figure S4), supporting the idea that they are indeed bound by CSR-1. Interestingly, similar conclusions can be drawn from HRDE-1 (Shirayama et al., 2012) and WAGO-1 (Gu et al., 2009) IP data (Fig S4): these Argonautes contain significant amounts of 22G-RNAs from typical CSR-1 target genes, and their uridylation is lower than is normally seen in CSR-1 IPs (Figure S4). Everything considered, we conclude that annotated CSR-1 targets are not exclusively targeted by CSR-1, but also by other Argonaute proteins like WAGO-1 and HRDE-1. Likewise, typical HRDE-1 or WAGO-1 targets are also targeted by CSR-1.

These insights might well relate to the above described result that in L1 larva, the relative abundance of 22G RNAs from typical CSR-1 targets tends to be reduced in mutator, or *prg-1* mutants (Figure 4B). This drop may reflect a pool of non-CSR-1 bound 22G-RNAs coming from typical CSR-1 targets that is present in wild-type animals, but absent in mutator mutants. Consistent with this idea, uridylation frequencies of 'CSR-1-pathway' 22G-RNAs are higher in *prg-1*/mutator pathway mutants (Figure 4C), potentially because a non-CSR-1-bound, and thus a non-uridylated pool of these 22G-RNAs is lost in these mutants. This strongly

suggests that a significant fraction 'CSR-1'-22G RNAs is in fact made through the *prg-1*/mutator pathway, and is loaded into non-CSR-1 Argonaute proteins. We hypothesized that the mutator-induced sterility phenotype is triggered by increased mutator activity on CSR-1 target genes, which to a large extent represent genes required form germ cell development, and that this results in a disturbed balance between CSR-1-mediated gene activation and HRDE-1-mediated silencing.

HRDE-1 is required to trigger mutator-induced sterility.

We next asked how CSR-1-target derived 22G-RNA pools change when animals develop from L1 to L2 larva. Transcription is activated in the germ cells during L1 stage, accompanied by proliferation, which continues in L2 and later stages (Kimble and Crittenden, 2005). In absence of straight IP-experiments, that will be hard to perform on limited numbers of young larva, we used again uridylation frequencies of 22G-RNA pools as an indicator for their physical presence in either CSR-1 (high uridylation) or in other Argonautes (low uridylation). In wild-type animals, uridylation frequencies of typical CSR-1-22G RNAs increase when animals develop from L1 into L2 larva, and at the same time the abundance of these 22G-RNAs drops strongly (Figure 4B,C). Given the onset of transcription during L1 stage, these two observations are possibly related through a mechanism known as target-dependent trimming and tailing of small RNA molecules (Ameres et al., 2010). In contrast, a significant decrease in uridylation of these 22G-RNAs is observed in animals that display mutator-induced sterility, as the animals progress from L1 to L2 stage, accompanied by only a small drop in 22G-RNA abundance (Figure 4C). This could point to excessive mutator activity acting on CSR-1 targets and loading of the resulting 22G-RNAs into non-CSR-1 Argonaute proteins. In effect, this would result in the replacement of a CSR-1-bound 22G-RNA population, by a WAGO-bound 22G-RNA population, both targeting germ cell-expressed genes. We hypothesize that this results in the inappropriate silencing of these genes and the observed germ cell defects.

Since these observations are rather indirect, and other scenarios might explain our primary observations regarding 22G-RNA abundance and uridylation, we sought for more direct evidence that an increase of gene-silencing activities is responsible for the mutator-induced sterility phenotype. The nuclear protein HRDE-1 is a good candidate to be loaded with CSR-1-target-derived 22G-RNAs if mutator activity on

these targets would increase, as HRDE-1 has been shown to be downstream of mutator proteins and PRG-1 (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). According to this model, mutator-induced sterility should at least be partially rescued by mutation of *hrde-1*. Indeed, when we perform a cross that normally results in mutator-induced sterility in absence of HRDE-1, the sterility phenotype is rescued to a large extent (Figure 5).

DISCUSSION

Maternal PRG-1 is required for initiation of transposon silencing

We describe that in *C. elegans*, piRNAs help to initiate a mutator-dependent response to transposons. In absence of parental memory of transposon silencing, animals require PRG-1 to initiate the biogenesis of a 22G-RNA pool that targets transposons. Interestingly, we demonstrate that the piRNAs present in the primordial germ cells are maternally derived, emphasizing the importance of maternal loading of piRNAs into the next generation. Finally, we were able to demonstrate that PRG-1 and the Argonaute protein HRDE-1 co-operate to silence transposition of *Tc1*. These findings extend the functional parallel between the C. elegans PRG-1 pathway and piRNA-activity in other animal species, including the importance of the maternal piRNA pool (Le Thomas et al., 2014a, 2014b). Our results indicate that maternal piRNAs are important for at least two reasons. One is the just mentioned initiation of transposon silencing. The second is that they help to keep mutator activity restricted to the proper targets. In absence of maternal piRNAs the chance of deleterious off-target silencing effects through the mutator genes significantly increases.

Parental memory of RNAe

Another force that keeps mutator activity on-target is the parental memory of the mutator genes themselves. Interestingly, both paternal as well as maternal inheritance of silencing information has been described before (Alcazar et al., 2008; Ashe et al., 2012; Grishok et al., 2000; Gu et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). In fact, the paternal inheritance is very strong in case of dsRNA-induced RNAi (Alcazar et al., 2008), whereas in case of RNAe, maternal inheritance tends to be dominant (Luteijn et al., 2012). It is currently unclear what determines

whether inheritance via the male or the female is strongest. It is also not clear how mutator memory is precisely transmitted. It seems most likely that this occurs in the form of 22G-loaded Argonaute proteins, such as HRDE-1 (Buckley et al., 2012), but how these Argonaute proteins drive the renewed synthesis of 22G-RNAs from their targets is not known.

Molecular mechanism behind Mutator-induced sterility

Our results suggest that inappropriate targeting of CSR-1 target genes by mutator complexes leads to sterility. Our data indicate that this occurs in wild-type animals as well, but that during larval development the non-CSR-1 bound 22G-RNAs from these targets drop in relative abundance, while they increase when mutator activity is re-started in absence of PRG-1. Of course, absolute levels of 22G-RNA populations are impossible to determine using small RNA sequencing, since a strong change in any other abundant small-RNA pathway would directly influence apparent abundance of all other small-RNA pathways. Hence, the trends we present should be interpreted with care. Nevertheless, all observed changes are consistent with a model in which, under normal circumstances, there is a low level of mutator activity acting on CSR-1 targets. This is not resulting in silencing, since there is sufficient CSR-1 present to counteract silencing activities. Upon activation of mutator activity in absence of both maternal piRNAs and parental 22G-memory, mutator activity increases on the CSR-1 targets, increasing the silencing signal on these genes to an extent that CSR-1 cannot fully prevent silencing any longer.

CSR-1: Gene activation, or protection from silencing?

How CSR-1 counteracts silencing has not been fully resolved. This Argonaute protein has been shown to act on chromatin (Claycomb et al., 2009; Wedeles et al., 2013), and hence a role for CSR-1 in maintaining open chromatin in a more-or-less direct manner is possible. However, we note another possibility, inspired by our finding that in wild-type animals, CSR-1-bound 22G-RNAs drop abruptly when animals develop from L1 to L2 stage (Figure 4B). This is accompanied by a rise in uridylation frequency, and occurs at a time when the PGCs become transcriptionally more active. Potentially, CSR-1 could act to directly destabilize 22G-RNAs in a target-dependent manner, similar to what has been described for miRNAs (Ameres et al.,

2010). This could serve as a threshold mechanism, ensuring that a minimum level of 22G-RNA production needs to be achieved before sufficient 22G-RNAs find their way into silencing-mediating Argonaute proteins and silencing takes effect. Mutator activity could have evolved for that very purpose: to overcome CSR-1-mediated clearance of 22G-RNAs, by boosting RdRP activity. We note that the existence of RdRP-containing mutator-foci (Phillips et al., 2012) is perfectly compatible with this idea.

ACKNOWLEDGEMENTS

We thank Emil Karaulanov for help in bioinformatics analysis and Chung-Ting Han for help in generating small RNA libraries and sequencing. M.P. was supported by a Boehringer Ingelheim Fonds PhD Fellowship. This work was further supported by a Deutsche Forschungsgemeinschaft grant KE 1888/1-1 (Project Funding Programme; R.K.) and a grant from Fundação para a Ciência e Tecnologia (FCT) SFRH/BD/51001/2010 (B.A.).

AUTHOR CONTRIBUTIONS

B.A. designed, executed and interpreted experiments, and performed computational analysis. M.P. planned, executed and interpreted experiments. R.F.K. designed the study, interpreted results and wrote the manuscript.

MATERIALS AND METHODS

Sample preparation

C. elegans L1 larvae were obtained by bleaching gravid adults and letting the eggs hatch in M9 over-night. L1 cross-offspring larvae were obtained single picking 200 eggs to an unseeded NGM plate, bleaching those eggs to remove any bacteria that was carried along, and allowed to hatch over-night. L1 larvae were then re-suspended in M9. To identify cross offspring, males were carrying the *punc-119::GFP* transgene.

L2 larvae were single picked and washed in M9 buffer for each sample. Cross offspring was identified by the presence of a marker carried only by the male.

Total RNA isolation

150 L1 or 50 L2 C. elegans larvae were washed in M9 buffer (22nM Na2HPO4, 33 mM KH2PO4, 86 mM NaCl, 1 mM MgSO4) and digested in lysis buffer (200mM NaCl, 100 mM Tris pH 8.5, 50 mM EDTA, 0.5 % SDS, 200 ug/mL Prot-K) for 3h at 65°C followed by 15 min at 95° C to denaturate the Prot-K. Lysate was then incubated with DNase I (NEB) for 30 min at 37° C. Total RNA was then isolated using TRIZOL-LS using the manufacturer instructions and dissolved in 8uL of H2O.

Library preparation

All 8uL of total RNA was treated with 5 U of tobacco acid phosphatase (Epicenter) at 37°C for 2 h to digest 5' tri- and di-phosphates to mono-phosphates. RNA was size-selected between 15- to 35-nt on 15% TBE-urea gel. Gel-purified RNA was eluted overnight in 300 mM NaCl and then precipitated with 100% isopropanol and Glycoblue for 1 h at -20°C. The pellet was washed once with 75% ethanol and dissolved in nuclease-free water. Then, this purified fraction was confirmed by Bioanalyzer Small RNA chip (Agilent). Library preparation was based on the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England BioLabs) with slight modification. In brief, small RNA was first ligated to the 3' adapter and then the 5' adapter, both of which contained 4 random bases and were chemically synthesized by Bioo Scientific. Adapter-ligated RNA was reverse-transcribed and PCR-amplified for 14 cycles using index primers. The PCR-amplified cDNA construct was purified using AMPure XP beads (Beckman Coulter). The purified PCR reaction was checked on the Bioanalyzer using High Sensitivity DNA chip (Agilent). Size selection of the small RNA library was done on LabChip XT instrument (Perkin Elmer) using DNA 300 assay kit. Only the fraction containing 140-165 bp was pooled in equal molar ratio. The resulting 10 nM pool was denatured and diluted to 10 pM with 5% PhiX spike-in and sequenced as single-read on HiSeq 2500 (Illumina) in rapid mode for 50 cycles using on-board cluster generation. Sample "L1 N2" was sequenced as single-read on Miseq (Illumina). After demultiplexing, on average 35 million passing filter reads were obtained per sample.

Data analysis

The raw reads in FastQ format were filtered from 3' adapter sequences and size-selected in the range 15-35 bases (plus 8 bases random barcodes) using cutadapt v.1.2.1 (Martin, 2011) using parameters: -a AGATCGGAAGAGCACACGTCT -O 8 -m 23 -M 43 . Subsequently, PCR clonal reads were deduplicated using Bash and Awk commands. All reads containing low-quality (Phred+33 score less than 20) bases were filtered with the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/; fastq_quality_filter -q 20 -p 100 -Q 33), then the files were reformatted from FastQ into tabular format, sorted and deduplicated based on full sequence identity (library insert plus 5' and 3' random barcodes of 4 nucleotides), and finally converted back to FastQ format for mapping. Quality assessment of the raw and processed data was done with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Mapping to the C. elegans genome reference WS244 was performed using Bowtie v1 (Langmead et al., 2009) with parameters: -v0 -M1 --best --strata --nomaqround --tryhard --trimm5 4 --trim3 4. Unmapped reads where then filtered for the ones ending with a "T" using custom python scripts and remapped using bowtie using the same parameter but trimming one extra 3' base. These newly mapped reads were considered mono-uridylated. This filtering/mapping cycle was performed 10 times in order to map up to penta-urydilated reads. All mapped reads where then annotated using bedtools intersect (Quinlan and Hall, 2010) with a customized WS244.gff3 (ftp://ftp.wormbase.org/pub/wormbase/releases/WS244/species/c_elegans/PRJNA13758/c_elegans/PRJNA13758.WS244.annotations.gff3.gz); using the following parameters: -abam -b[custom annotated gff3] -bed -wa -wb. Mapped and annotated reads were subsequently filtered for size and starting nucleotide using custom made python scripts and normalized to total of non-structural reads between 18 and 30 nucleotides. Structural reads were considered as reads that mapped rRNAs, tRNAs and snoRNAs. During the analysis we considered miRNA reads that were 22 to 24 nt long sense to annotated miRNAs, we considered piRNA reads that were 21 nt long, started with a "T" and were sense to annotated piRNAs, and 22G reads that were 20 to 23 nt long, started with a "G" and map antisense to genes, transposable elements or pseudogenes. In order to categorize genes, we retrieved WAGO-1 targets, ERGO-1 targets, mutator targets, ALG-3/4 targets and CSR-1 targets from (Claycomb et al., 2009; Conine et al., 2010; Gu et al., 2009; Phillips et al., 2014; Vasale et al., 2010) respectively. For each protein coding gene saw if it

was a WAGO-1 target, if not we saw if it was an ERGO-1 target, if not we saw if it was a mutator target and so on, in the previously mention order. The validation of a gene in a category would automatically exclude it from the following categories; this was to insure that there would be no duplication, since some categories overlap partially.

WAGOs IP data analysis

Sequencing data from CSR-1 IP, WAGO-1 IP and WAGO-9 IP were obtained from (Claycomb et al., 2009; Gu et al., 2009; Shirayama et al., 2012) respectively. The raw reads in FastQ format were filtered from 5' barcodes 3' adapter sequences using a custom python script, mapped and processed to the C. elegans genome reference WS224 as mentioned before with the exception that any gene was allowed to be in several "target categories".

Statistical analysis

Samples were compared using 2-tailted t-test, assuming a normal distribution. L1 samples OH441(2x) and N2 (1x) were considered wild-type triplicates.

Immunohistochemistry

Adult worm gonads were dissected in a poli-L lysine (Sigma P0425) and freeze cracked. Worms were then fixed for 2 minutes in ice cold methanol followed by 2 minute in acetone, and washed 3 times 10 min with PBS-tween (0.05%). Blocking was done with 10% lamb serum in PBT-Tween20 (0.05%). Samples where then incubated O.N. at 4° C with primary antibody (mouse anti-PGL-1 (Capowski et al., 1991) in block buffer). After O.N. incubation, samples were washed 3x 10 min with PBS-Tween20 (0.05%), incubated with secondary antibody in block buffer for 1h at room temperature, washed 3x 10 min with PBS-Tween20 (0.05) and mounted Fluoroshield with DAPI (Sigma F6057) Samples were then visualized using a Leica TCS SPE confocal.

REFERENCES

De Albuquerque, B.F.M., Luteijn, M.J., Cordeiro Rodrigues, R.J., van Bergeijk, P., Waaijers, S., Kaaij, L.J.T., Klein, H., Boxem, M., and Ketting, R.F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28, 683–688.

Alcazar, R.M., Lin, R., and Fire, A.Z. (2008). Transmission dynamics of heritable silencing induced by double-stranded RNA in Caenorhabditis elegans. Genetics 180, 1275–1288.

Ameres, S.L., Horwich, M.D., Hung, J., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science 328, 1534–1539.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150, 88–99.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M.E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E. a (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science (80-.). 337, 574–578.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31, 67–78.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489, 447–451.

Capowski, E.E., Martin, P., Garvin, C., and Strome, S. (1991). Identification of grandchildless loci whose products are required for normal germ-line development in the nematode Caenorhabditis elegans. Genetics 129, 1061–1072.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van Wolfswinkel, J.C., Chaves, D. a, Shirayama, M., Mitani, S., Ketting, R.F., et al. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139, 123–134.

Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 107, 3588–3593.

Conine, C.C., Moresco, J.J., Gu, W., Shirayama, M., Conte, D., Yates, J.R., and Mello, C.C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell.

Cox, D.N., Chao, a, Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. 12, 3715–3727.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell 31, 79–90.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94–108.

Grishok, A., Tabara, H., and Mello, C.C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science 287, 2494–2497.

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J., et al. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell 36, 231–244.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151, 1488–1500.

Houwing, S., Berezikov, E., and Ketting, R.F. (2008). Zili is required for germ cell differentiation and meiosis in zebrafish. EMBO J. 27, 2702–2711.

Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. Cell 94, 635–645.

Ketting, R.F. (2011). The many faces of RNAi. Dev. Cell 20, 148–161.

Ketting, R.F., Haverkamp, T.H. a, Van Luenen, H.G. a M., and Plasterk, R.H. a (1999). mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of werner syndrome helicase and RNaseD. Cell 99, 133–141.

Kimble, J., and Crittenden, S.L. (2005). Germline proliferation and its control. WormBook 1–14.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150, 78–87.

Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO J. 31, 3422–3430.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell 136, 656–668.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10.

Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the C . elegans germline. 1433–1444.

Phillips, C.M., Montgomery, B.E., Breen, P.C., Roovers, E.F., Rim, Y.-S., Ohsumi, T.K., Newman, M.A., van Wolfswinkel, J.C., Ketting, R.F., Ruvkun, G., et al. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. Curr. Biol. 24, 839–844.

Qiao, L., Lissemore, J.L., Shu, P., Smardon, A., Gelber, M.B., and Maine, E.M. (1995). Enhancers of glp-1, a gene required for cell-signaling in Caenorhabditis elegans, define a set of genes required for germline development. Genetics 141, 551–569.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Dev. Cell 27, 656–663.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150, 65–77.

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L.D., Mitchell, J., Sakaguchi, A., Miska, E. a, and Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to caenorhabditis elegans Piwi mutants. Cell Rep. 7, 762–773.

Stoeckius, M., Grün, D., and Rajewsky, N. (2014). Paternal RNA contributions in the Caenorhabditis elegans zygote. EMBO J. 33, 1740–1750.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99, 123–132.

Le Thomas, A., Marinov, G.K., and Aravin, A. a (2014a). A transgenerational process defines piRNA biogenesis in Drosophila virilis. Cell Rep. 8, 1617–1623.

Le Thomas, A., Stuwe, E., Li, S., Du, J., Marinov, G., Rozhkov, N., Chen, Y.-C.A., Luo, Y., Sachidanandam, R., Toth, K.F., et al. (2014b). Transgenerationally inherited piRNAs trigger piRNA

biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev. 28, 1667–1680.

Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U. S. A. 107, 3582–3587.

Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. Curr. Biol. 18, 861–867.

Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013). Protection of Germline Gene Expression by C. elegans Argonaute CSR-1. Dev. Cell 27, 664–671.

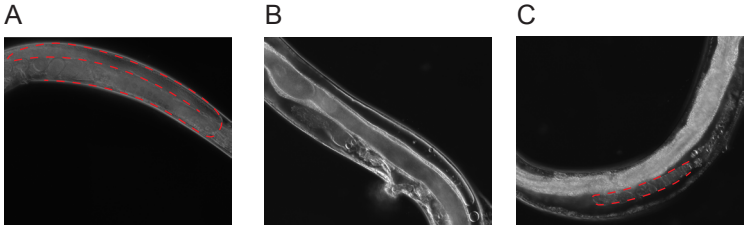
Van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell 139, 135–148.

Zhang, C., Montgomery, T. a, Gabel, H.W., Fischer, S.E.J., Phillips, C.M., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 108, 1201–1208.

SUPLEMENTAL INFORMATION

Supplemental table 1. General statistics of small RNA sequencing libraries.

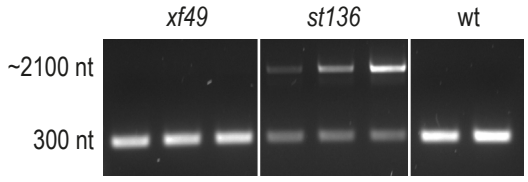
	strain	total mapped reads	18nt - 30nt reads	Non-structrural reads
L2 larvae	N2	2325113	1590416	361940
	N2	740650	506737	79221
	RFK179	993562	614892	106718
	RFK179	1962593	1238411	192970
	RFK232	1152744	767006	165669
	RFK232	1206209	790264	153039
	RFK245♀x RFK179♂	1928221	1243183	343456
	RFK245♀x RFK179♂	2729489	1752633	326982
	RFK231♂ x RFK232♀	1924652	1261777	260325
	RFK231♂ x RFK232♀	1816006	1075126	203344
L1 larvae	OH441	1529601	1158242	339308
	OH441	2631423	2008094	495152
	RFK240	2913868	2135282	525812
	RFK240	3786140	2577927	569112
	RFK231	293937	187826	47302
	RFK231	1178626	818640	198765
	RFK231	2395818	1803061	404776
	SX922	1551145	1021268	224469
	SX922	2803688	1961449	471299
	SX922	2878124	1658862	405842
	N2 ♂ x SX922♀	566551	336655	123145
	N2 ♂ x SX922♀	1268714	957918	285274
	N2 ♂ x SX922♀	1464043	993611	233270
	RFK231♂ x RFK240♀	2145206	1502899	448209
	RFK231♂ x RFK240♀	2965104	1773575	344584
	RFK231♂ x RFK240♀	824708	562172	160774
	RFK231♂ x RFK232♀	192668	115722	24414
	RFK231♂ x RFK232♀	1819925	1112448	232654
	N2	647601	519557	105177



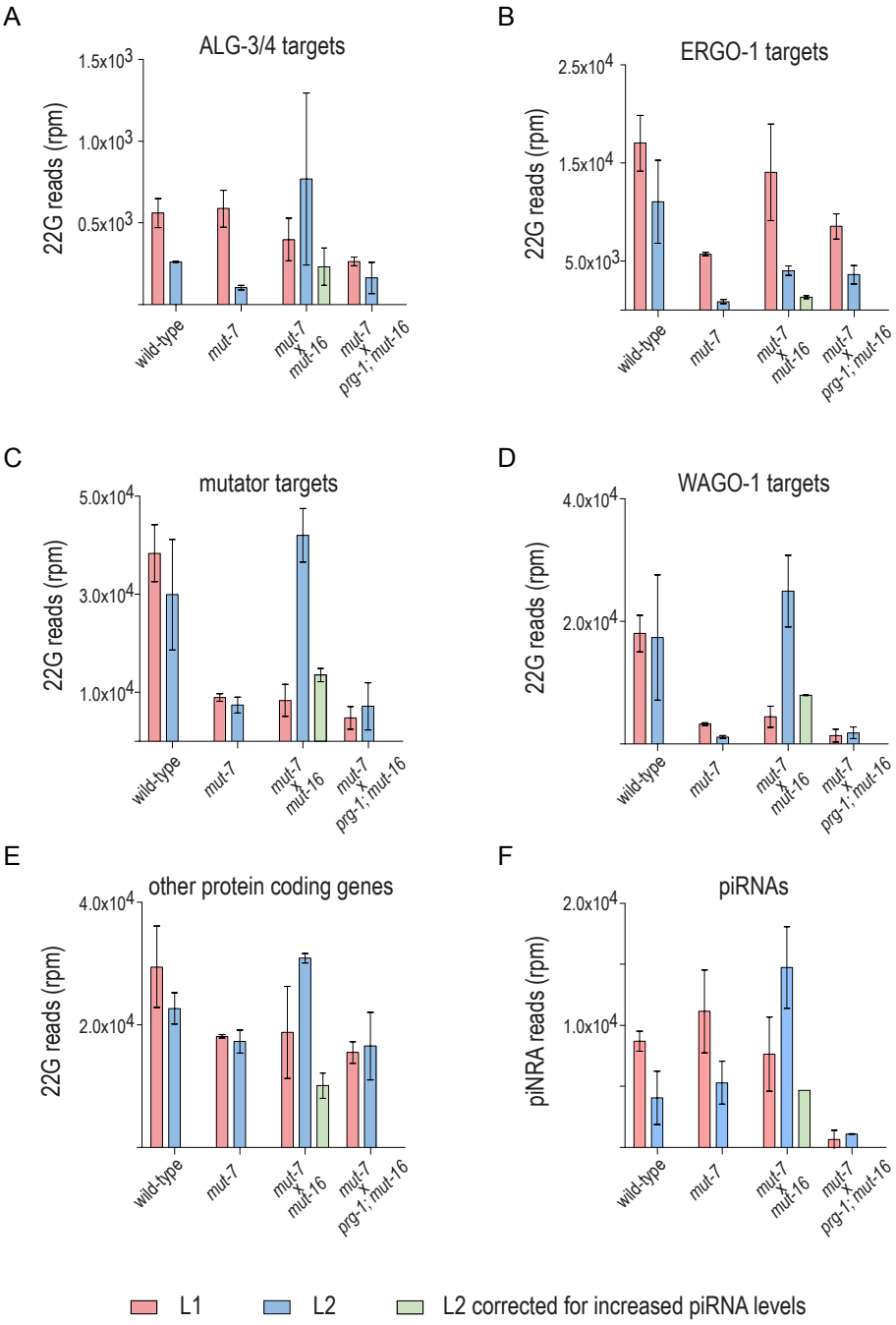
Supplemental figure 1. Activation of the mutator pathway in the absence of PRG-1 results in germlineless worms. DIC picture of **A)** wild-type worm and **B,C)** *prg-1(n4357);mut-7(+/pk204); mut-16(+/pk710)* F1 from a cross between *mut-16(pk710); prg-1(n4357)* ♀ and *mut-7(pk204); prg-1(n4357)* ♂. The germline of each worm is outlined withe the dashed line.

Supplemental table 2. List of C. elegans strains used.

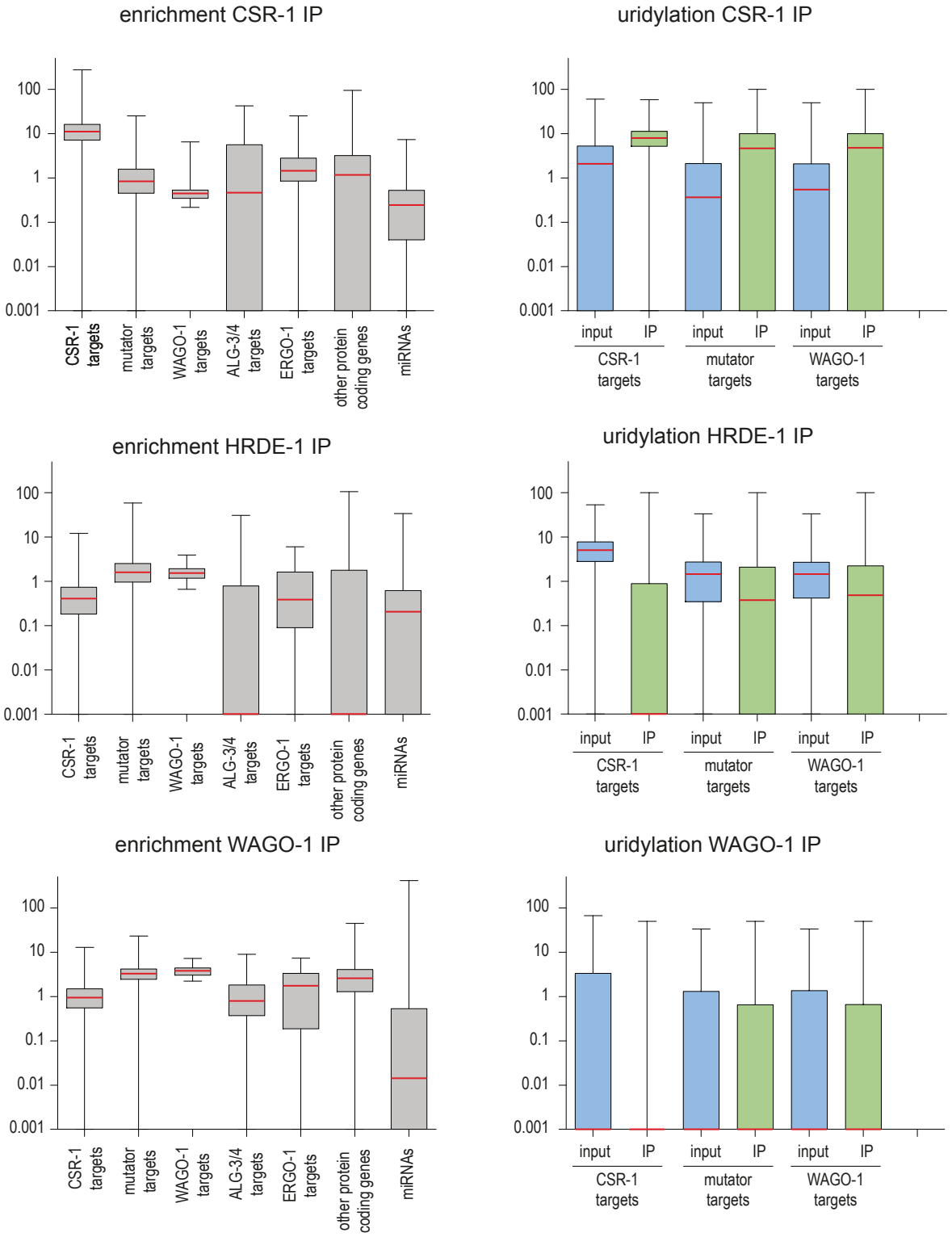
STRAIN	GENOTYPE
OH441	*otIs45 [Punc119::GFP] V*
RFK162	*hrde-1(tm1200) III*
RFK179	*mut-7(pk204) III*
RFK231	*mut-7(pk204) III; mjSi22 [Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; otIs45 [Punc119::GFP] V*
RFK232	*mut-16(pk710) I;prg-1(n4357) I*
RFK233	*mut-16(pk710) I;prg-1(n4357) I; unc-22(st136::Tc1) IV*
RFK234	*mut-7(pk204) III; pid-1(xf35)II*
RFK235	*hrde-1(tm1200) III;mut-7(pk204) III; pid-1(xf35) II*
RFK236	*hrde-1(tm1200) III; mut-16(pk710) I; prg-1(n4357) I*
RFK237	*mut-7(pk204) III; prg-1(n4357) I*
RFK238	*mut-14 (mg464) V; smut-1 (tm1301) V;mut-16 (pk710) I; prg-1(n4357) I*
RFK239	*mut-14 (mg464) V; smut-1 (tm1301) V;mut-7(pk204) III;prg-1(n4357) I*
RFK240	*mut-16(pk710) I*
RFK241	*hrde-1(tm1200) III;mut-16(pk710)I*
RFK242	*prg-1(n4357) I; mjIs144 [Pmex-5::egfp::his-58::21UR-1_as::tbb-2(3'UTR)] [RNAe] II*
RFK243	*mut-14 (mg464)V; smut-1 (tm1301) V;prg-1(n4357) I; mjIs144 [Pmex-5::egfp::his-58::21UR-1_as::tbb-2(3'UTR)] II*
RFK244	*mut-14 (mg464)V; smut-1 (tm1301) V; prg-1(n4357) I*
SX922	*prg-1(n4357) I*
RFK245	*mut-16(pk710) I; mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I*
RFK246	*prg-1(n4357) I; hrde-1(tm1200) III*
RFK247	*prg-1(n4357) I; unc-22(st136::Tc1) IV*
RFK248	*hrde-1(tm1200) III; unc-22(xf49) IV*
RFK249	*prg-1(n4357) I; hrde-1(tm1200) III; unc-22(xf49) IV*
NL3643	*unc-22(st136::Tc1) IV*



Supplemental figure 2. HRDE-1 and PRG act redundantly to silence Tc1 transposons. PCR amplification of the *Tc1* insertion at the *st136* locus in reverted *hrde-1(tm1200);prg-1(n3457);unc-22::Tc1(st126)* animals.



Supplemental figure 3. Column chart showing antisense 22G RNAs to **A)** ALG-3/4 target genes, **B)** ERGO-1 target genes, **C)** mutator target genes, **D)** WAGO-1 target genes, **E)** remaining protein coding genes. F) levels of piRNAs. Values are in "reads per million" of non-structural reads (rpm) and errors represent standard deviation between at least two biological replicates. In the L2 larvae offspring of mut-7 animals crossed with mut-16 animals we observed an overall increase of small RNAs which we attribute to the difficulty to have a synchronized population, to overcome that used piRNA values to correct for the increase in germ cell number.



Supplemental figure 4. Enrichment and uridylation levels for 22G siRNAs antisense to protein coding genes that co-immunoprecipitate with CSR-1, HRDE-1 and WAGO-1. 22G siRNAs targeting "CSR-1 targets" are not enriched in HRDE-1 and WAGO-9, but we believe that is an artefact due to normalization to total number of non-structural reads. When compared to miRNAs depletion (a small RNA class known to not interact with HRDE-1 or WAGO-9, the levels of these 22G siRNA are higher (less depleted) suggesting that they are interacting with HRDE-1 and WAGO-1 but at a low level. Moreover uridylation levels of those 22G siRNAs are highly reduced when compared to input levels, further reinforcing a possible interaction with HRDE-1 and WAGO-1.

Chapter 4

pid-2 is a novel factor involved in piRNA induced silencing that acts downstream mutator activity

Bruno F.M. de Albuquerque, Maartje J. Luteijn, Petra van Bergeijk and René F. Ketting

SUMMARY:

One of the challenges faced by germ cells is to repress molecular parasites, such as transposons and retro-viruses that when left unchecked can lead to fertility problems. Organisms have developed several mechanisms to accomplish this task, such as the use of Piwi-interacting RNAs (piRNAs) to identify and silence foreign sequences. In *C. elegans*, the PIWI Related Gene-1 (PRG-1) uses a repertoire of more than 30000 piRNAs to recognize foreign sequences and trigger transcriptional gene silencing. Through a yet unknown mechanism, silencing can become independent of PRG-1. This state, named as RNAe, is heritable and dependent on mutator proteins and the Argonaute protein HRDE-1. Using a strain carrying a piRNA sensor we performed a forward mutagenesis screen for piRNA-Induced-Silencing-Defective worms and we isolated *pid-2*. Worms deficient for *pid-2* show secondary siRNA response triggered by PRG-1 but fail to trigger RNAe at the piRNA sensor. Moreover, removal of PID-2 does not re-activates a piRNA in RNAe state. Taken together, our results suggests that *pid-2* is linking secondary siRNA activity and RNAe.

INTRODUCTION

An essential property of an organism is to spread its genetic material through countless generations. It is of utmost importance to preserve genome stability of germ cells and in order to do so, organism evolved several mechanisms. One of such mechanisms is the use of PIWI Argonaute proteins and their piRNA co-factors to silence foreign molecular species such as transposons and retro-viruses. Absence of a functional piRNA pathway often leads to defects in germline development that result in sterility (Ghildiyal and Zamore, 2009; Luteijn and Ketting, 2013; Malone and Hannon, 2009).

Similar to other organisms, the nematode *C. elegans,* uses the PIWI Related Gene-1 (PRG-1) and its associated piRNAs to target foreign transcripts (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). *C. elegans* piRNAs, also called 21Us, are individually transcribed from discrete genetic units in the form of 28-35 nucleotide long precursors that are further processed into the mature 21U piRNA. There are more than 30000 identified piRNAs, grouped into 2 classes that differ in how piRNA precursors are transcribed (Gu et al., 2012). The most abundant ones (Type I) share a consensus motif upstream each locus (Ruby et al., 2006) and their transcription depends on SNP-3 and PRDE-1 (Kasper et al., 2014; Weick and Miska, 2014). Type II piRNAs are less expressed and thought to be a product of early Pol II termination during gene transcription (Gu et al., 2012). In both types, piRNA precursors contain a 5' 7-methylguanylate cap and transcription starts precisely 2 nucleotides upstream of the 5' U of the mature piRNA, at the YRNT motif, and extends up to 15 bases after the 3' end of the mature piRNA (de Albuquerque et al., 2014; Gu et al., 2012; Weick et al., 2014). After transcription piRNA precursors are processed into a mature 21 nucleotide long piRNA. How exactly piRNA precursor are processed is still unknown, but recently studies found new factors involved in this process, namely PID-1 and the TOFU genes (de Albuquerque et al., 2014; Goh et al., 2014). The last step in piRNA biogenesis is 3' 2-O-methylation by HENN-1 (Kamminga et al., 2012).

After biogenesis, piRNAs are loaded onto PRG-1 where they are used as co-factors to identify target transcripts through incomplete base-pair complementarity (Bagijn et al., 2012; Lee et al., 2012). PRG-1 has a catalytically active PIWI domain; however it doesn't appear to cleave targeted transcripts. Instead, upon target recognition, PRG-1 triggers the production of secondary 22G siRNAs near the

targeted site (Bagijn et al., 2012; Lee et al., 2012). This 22G siRNA production is dependent on mutator genes, and the Argonaute proteins WAGO-1 and HRDE-1. By an as yet elusive mechanism, mutator activity can in turn trigger a nuclear RNAi response that becomes independent on PRG-1. This state, named as RNAe, is heritable and is able to spread in trans to similar sequences (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012).

In this study we describe the identification of the gene *pid-2*, which we isolated in forward genetic a screen for factors involved in piRNA Induced silencing (de Albuquerque et al., 2014). PID-2 is a novel a factor that is necessary for piRNA induced silencing and acts downstream the mutator genes.

RESULTS AND DISCUSSION

Isolation of piRNA induced silencing defective mutants.

We previously described a screen to isolate mutants defective for piRNA induced silencing where we identified the novel factor *pid-1* and new alleles for *mut-7*, *rde-3* and *hrde-1* (de Albuquerque et al., 2014). In short, we mutagenized worms carrying a germline specific GFP::H2B transgene that contains a *21ur-1* binding site at the 3'end (Bagijn et al., 2012). In order to increase sensibility and be able to score mutants that mildly impair piRNA induced silencing we performed the screen in a *henn-1* background. Henn-1 is a methyl-transferase that catalyzes the last step of piRNA biogenesis, by 2'-O-methylating the 3' end of the 21U small RNA (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). The piRNA pathway in *henn-1* mutants is only mildly impaired and mutants only partially silence the piRNA sensor (Kamminga et al., 2012). From the screen we also isolated the allele *xf23* defining a locus we named *pid-2*.

pid-2 is Y48G1C.1

In order to identify *pid-2* we performed genome wide re-sequencing of the mutant strain carrying the *xf23* allele. We found 96 non-synonymous mutations (supplemental table 1) from which 5 were premature STOP mutations and 6 were in genes previously described to be enriched in primordial germ cells (PGC) (Spencer et al., 2011). Since factors involved in the biogenesis of piRNAs in *C. elegans* are

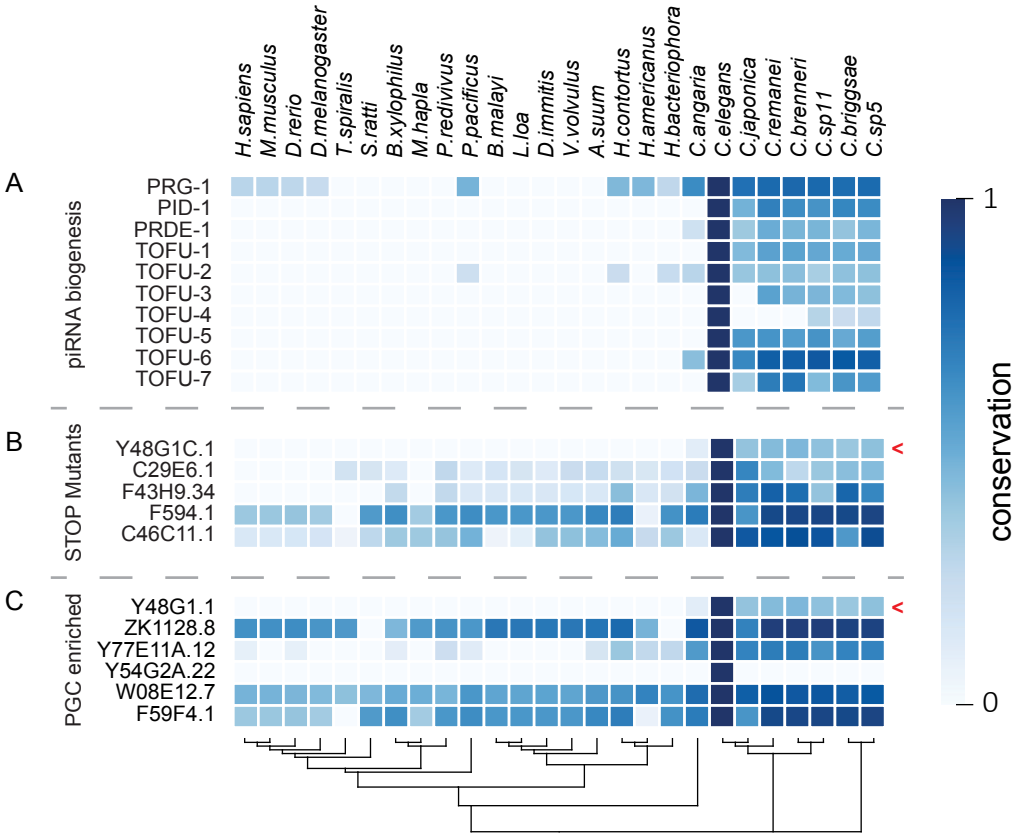


Figure 1. Y48G1C.1 is conserved among the Caernohabditis genera. Heatmap showing evolutionary conservation between several species of: **A)** factors involved in piRNA biogenesis, **B)** genes with STOP mutations present in the strain carrying *pid-2(xf23)*, **C)** genes with non-synonymous mutations in the strain carrying *pid-2(xf23)* that are enriched in PGC (Spencer et al., 2011).Conservation values were determined by first BLASTing each protein sequence against the proteome of each specie, then the bit score value of the best match was normalized to the bit score of the best match in C.elegans.Y48G1C.1 (red arrows <) show a similar profile to piRNA biogenesis factors, is enriched in PGC and contains a STOP mutation in the strain carrying the *pid-2(xf23)* allele.

mostly restricted to the *Caenorhabditis* genera (figure 1A), we asked if any of the genes that either contained a STOP mutation (figure 1B) or whose expression was enriched in PGCs (figure 1C), showed a similar conservation profile. We found that Y48G1C.1 contained a STOP mutation, is enriched in PGCs and is conserved among the Caenorhabditis genera only. To test if *pid-2* is Y48G1C.1 we crossed the deletion allele for Y48G1C.1(*tm1614*) with the piRNA sensor and we observed de-silencing of the piRNA sensor (figure 2A). Therefore we conclude that *pid-2* is Y48G1C.1.

PID-2 is a 454 amino-acid long protein without any known domains or motifs, it's conserved among the *Caenorhabditis* genera and does not have any apparent

homolog in other organisms. *pid-2(xf23)* leads to a premature STOP at position 122 (W122*) (figure 2B) whereas the deletion allele *tm1614* spans from upstream the start codon up to the end of the second exon.

pid-2 mutants partially silence the piRNA sensor and do not rescue RNAe.

PRG-1 mediated targeting of the piRNA sensor leads to nuclear RNAi response that becomes independent of PRG-1 (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). This state, referred to RNAe, is heritable and requires the Argonaute HRDE-1 and the mutator genes to be maintained. Interestingly, all known factors that are involved in the maintenance of RNAe are also required for "non-RNAe" silencing of the sensor. Therefore we asked if *pid-2* is able to rescue a sensor in the RNAe state. To investigate this we, crossed the *tm1416* allele with a strain mutant for *prg-1* that carried the 21U sensor in the RNAe state. We were unable to observe any reactivation of the sensor (figure 2A), either in the presence or absence of PRG-1. Therefore we conclude that *pid-2* is not involved in maintenance of RNAe.

We also noticed that in *pid-2* mutants the H2B::GFP intensity was seemingly lower than *prg-1* mutants. To address if the lower intensity is due to a lower transcript levels, we quantified sensor expression by RT-qPCR (figure 3). Indeed, both *pid-2* mutant alleles show a decrease in sensor transcripts when compared with *prg-1*

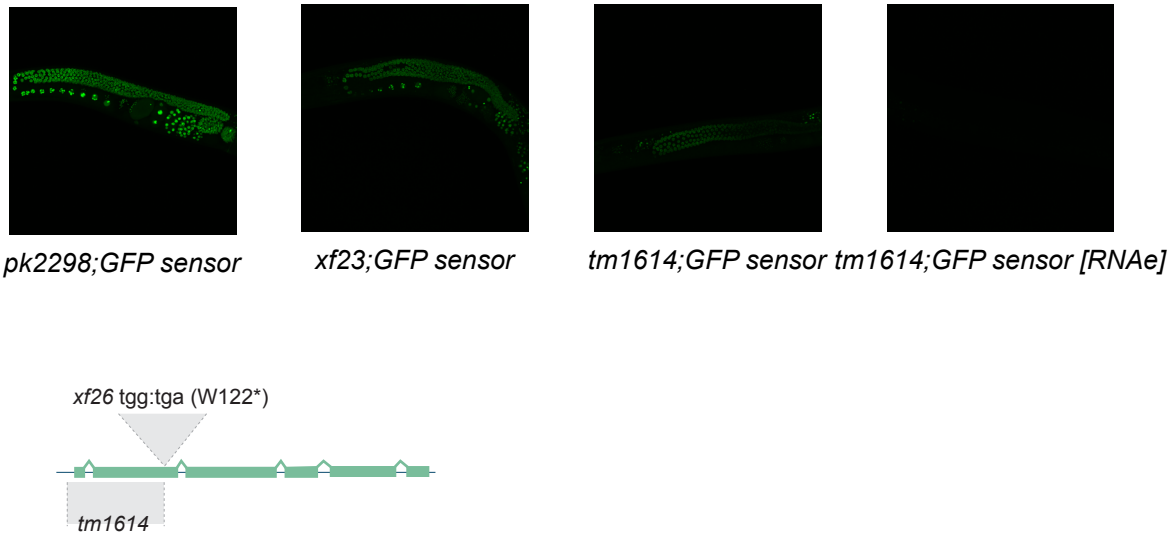


Figure 2. pid-2 is Y48G1C.1. A to D), confocal images of worms carrying the piRNA sensor in the specified mutant background. *pid-2(xf23)* and *pid-2(tm1614)* partially de-silence the piRNA sensor, and fail to rescue an sensor in the RNAe state. E) schematics of Y48G1C.1 and the effect of *xf23* and *tm1614*.

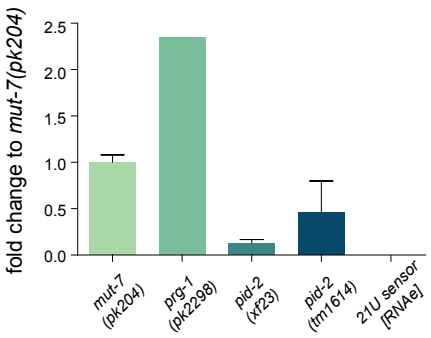


Figure 3. *pid-2* mutants show decreased expression of the piRNA sensor. Colum chart depicting mRNA levels of the piRNA sensor in the indicated genotypes. Gene expression values were determined by RT-qPCR and normalized to the transcript levels in the *mut-7(pk204)* mutant background. *pid-2* mutants expressing the sensor show a 4 fold depletion of sensor transcripts.

mutants or *mut-7* mutants. This correlates with the lower GFP signal observed and suggests that the sensor may still be partially silenced in *pid-2* mutants.

A similar partial silencing of the sensor is observed in *henn-1* mutants (Luteijn et al., 2012). Albeit in the absence of HENN-1, PRG-1 mediated recognition is not enough to elicit full silencing of the sensor, it is enough to trigger the onset of RNAe in approximately half of the worms (Luteijn et al., 2012). In contrast, we did not observe any stochastic onset of RNAe in *pid-2* mutants nor in *pid-2;henn-1* double mutants.

pid-2 mutants show normal levels of siRNAs

To determine the effect of *pid-2* on the siRNA populations we performed small RNA sequencing in young adult and gravid worms. In order to remove PCR duplications we used a random 4 nucleotide barcode in each adapter during library preparation. As expected, all samples show similar levels of miRNAs (figure 4B). Surprisingly we didn't see any significant change in piRNAs or 22Gs that could directly explain the activation of the sensor. piRNA levels are normal in the *tm1614* allele and possibly slightly elevated in the *xf23* allele (figure 4a) which exclude *pid-2* from piRNA biogenesis. However we do noticed a difference between both *pid-2* alleles. The *pid-2(xf23)* allele shows a 2 fold depletion of 22G secondary siRNA mapping anti-sense to transposons and pseudogenes (figure 4C and 4D). This depletion is also observed in the *prg-1(pk2298)* mutants. This may be explained by the age of the sequenced strain, as we kept both *prg-1(pk2298)* and *pid-2(xf23)* allele in continuous culture for more than a year whereas the *pid-2(tm1416)* strain we

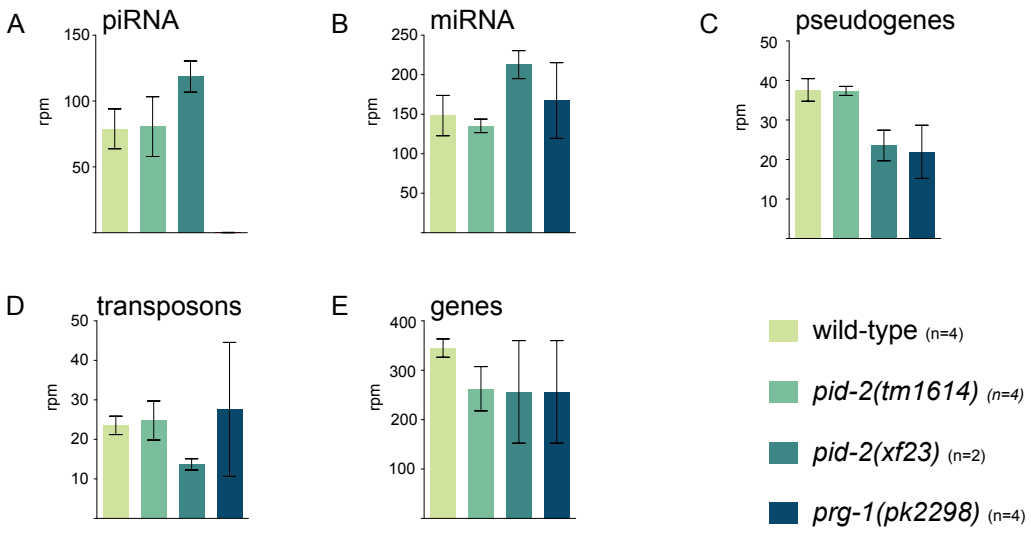


Figure 4. small RNA levels in pid-2 mutants are not severely affected. A) Column chart showing piRNA levels determined by small RNA-seq of total RNA from adult worms of the indicated genotype. piRNAs were considered as reads that were 20-21 nucleotide long, start with a U and were sense to annotated piRNA loci. Levels are indicated in “reads per million” of non-structural reads (rpm) and error bars represent standard deviation between biological replicates. **B)** Column chart showing miRNA levels determined by small RNA-seq of total RNA from aldult worms of the indicated genotype. miRNAs were considered as reads that were 22-24 nucleotide long and were sense to annotated miRNAs. Levels are indicated in “reads per million” of non-structural reads (rpm) and error bars represent standard deviation between biological replicates. **C to D)** Column chart showing 22G levels determined by small RNA-seq of total RNA from adult worms of the indicated genotype mapping antisense to pseudogenes, transposons and protein coding genes. 22G small RNAs were considered as reads that start with a G, are 20 to 23 nucleotides long and mapped antisense to annotated pseudogenes, transposable elements or genes respectively. Levels are indicated in “reads per million” of non-structural reads (rpm) and error bars represent standard deviation between biological replicates.

sequenced was heterozygous for *pid-2* one month prior to the sequencing. The drop in 22G levels observed may be due to the progressive loss of RNAe induced 22G. In the absence of PRG-1, RNAe is not renewed when lost and some transposons and pseudogenes may progressively loose RNAe silencing. If true, *pid-2(xf23)* may be phenocoping *prg-1* mutants.

With respect to 22G siRNAs that map anti-sense to protein coding genes, both alleles show a slight depletion of this class of small RNAs when compared to wild-type. Protein coding genes that are targeted by 22G siRNAs can be further divided into classes according to the factors required for 22G siRNA accumulation. Therefore we separated all protein coding genes into mutator targets, WAGO-1 targets, ERGO-1 targets, CSR-1 targets, ALG-3/4 targets. Mutator targets are those

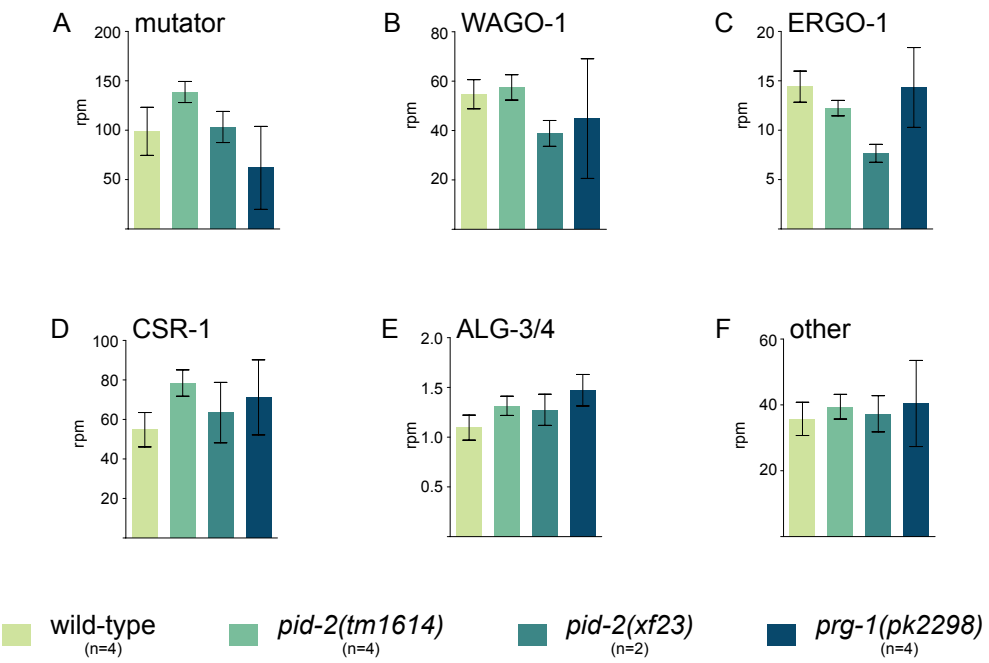


Figure 5. *pid-2* mutants show normal small RNA population against protein coding genes. Column charts showing 22G levels antisense to genes determined by small RNA-seq of total RNA from adult worms of the indicated genotype. Genes were separated into classes (more info in materials and methods) that represent the major siRNA classes in C.elegans. The class “others” represent genes that are not represented in any of the remaining classes. 22G small RNAs were considered as reads that start with a G, are 20 to 23 nucleotides long and mapped antisense to annotated pseudogenes, transposable elements or genes respectively. Levels are indicated in “reads per million” of non-structural reads (rpm) and error bars represent standard deviation between biological replicates.

who loose anti-sense 22G siRNAs in a *mut-16* or *mut-7* background (Phillips et al., 2014). WAGO-1 and ERGO-1 targets greatly overlap with mutator targets and were defined as genes whose anti-sense 22G siRNAs were enriched in a WAGO-1 immuno-precipitation or targeted by ERGO-1 bound 26G siRNAs respectively (Gu et al., 2009; Vasale et al., 2010). CSR-1 and ALG-3/4 targets do not overlap with mutator targets and were defined has genes whose anti-sense 22G siRNAs were enriched in CSR-1 immuno-precipitation or targeted by ALG-3/4 bound 26G siRNAs respectively (Claycomb et al., 2009; Conine et al., 2010). Protein coding genes that did not fit in any of the aforementioned classes were defined as “others”.

Again we didn’t see any big differences between wild-type and *pid-2* mutants that could easily explain the reactivation of the piRNA sensor. As expected, 22G siRNAs mapping antisense to CSR-1 targets, ALG-3/4 or genes that do not belong to any of the classes (others) are constant throughout all analyzed genotypes.

WAGO-1 is thought to be one of the Argonaute that acts downstream of PRG-1 recognition (Lee et al., 2012). As observed before for transposons and pseudogenes, we see a slight reduction of 22G mapping WAGO-1 targets in the *xf23 allele,* similar to what we observe in the PRG-1 mutant.

Interestingly, both *pid-2* mutants show a decrease in 22G siRNAs that map to ERGO-1 targets. Like in 22G siRNAs levels mapping transposons and pseudogenes, the reduction is stronger in the *pid-2(xf23)* allele, however, we did not observe any change in this class of 22G siRNAs in the *prg-1(pk2298)* mutant. ERGO-1 itself binds

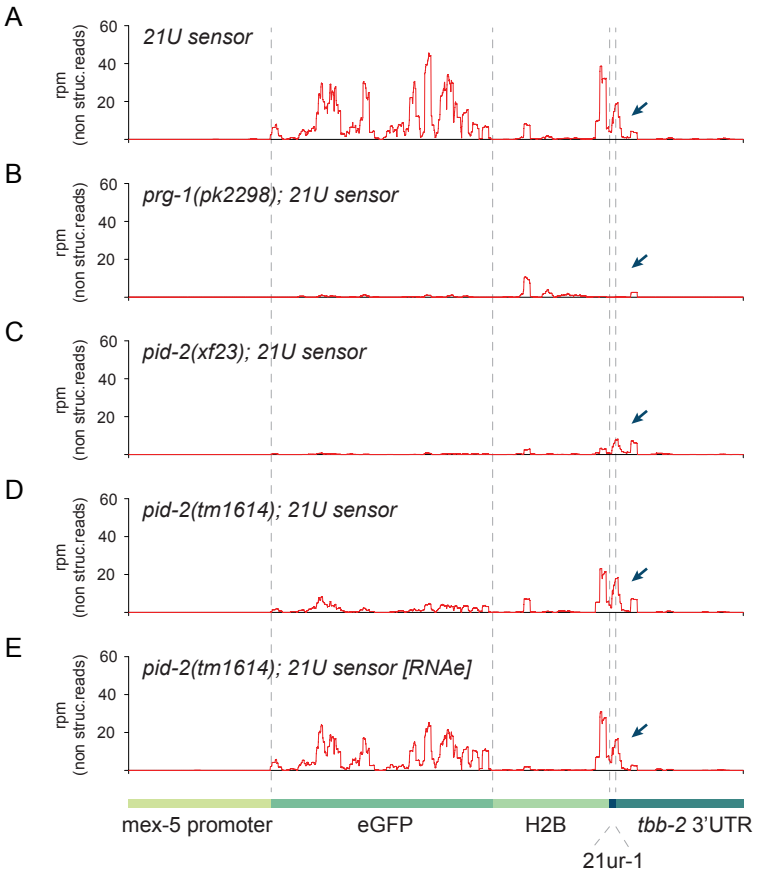


Figure 6. *pid-2* mutants show secondary siRNA response near the *21ur-1* recognition site. Plots show 22G coverage at the piRNA sensor determined by small RNA-seq of total RNA from adult worms with the indicated genotypes. 22G siRNAs were considered as reads that start with a G, are 20 to 23 nucleotides long and mapped antisense to the piRNA sensor. Levels are indicated in "reads per million" of non-structural reads (rpm). In wild-type worms, there are two cluster of 22G siRNAs, one near the *21ur-1* and other mapping the GFP. None of the strain show 22G siRNA mapping H2B. All 3 strains containing a *pid-2* mutation still show 22G siRNA production near the *21ur-1* recognition site, derived from mutator activity triggered by PRG-1 recognition (blue arrow).

to 26G siRNAs, but triggers 22G siRNAs production by the mutators which then stimulate nuclear RNAi mediated by the Argonaute NRDE-3 and the nuclear factors NRDE-1/2/4. Target recognition by ERGO-1 is thought to be made in early embryos which triggers nuclear RNAi silencing that is maintained in the soma by the NRDE pathway. This mechanism is very similar to what is observed with the establishment of the RNAe status triggered by PRG-1 induced silencing. In fact, *nrde-1*, *nrde-2* and *nrde-4* mutants are also able rescue the piRNA sensor from the RNAe (Ashe et al., 2012), which further suggest that both mechanism share the same molecular components. The fact that both *pid-2* alleles show a slight decrease in 22G siRNAs antisense to ERGO-1 targets, that otherwise are unaffected in the *prg-1(pk2298)*, coupled to the similar 22G siRNAs levels between *pid-2(xf23)* and *prg-1(pk2298)*, suggests that *pid-2* may act in a common step downstream the mutator activity.

pid-2 mutants show PRG-1 triggered secondary siRNA response against the piRNA sensor

We next looked at the 22G siRNA population anti-sense to the piRNA sensor in *pid-2* mutants. PRG-1 recognition triggers production of 22G siRNAs near 21U recognition sites (Lee et al., 2012). After the onset of RNAe, 22G siRNAs production

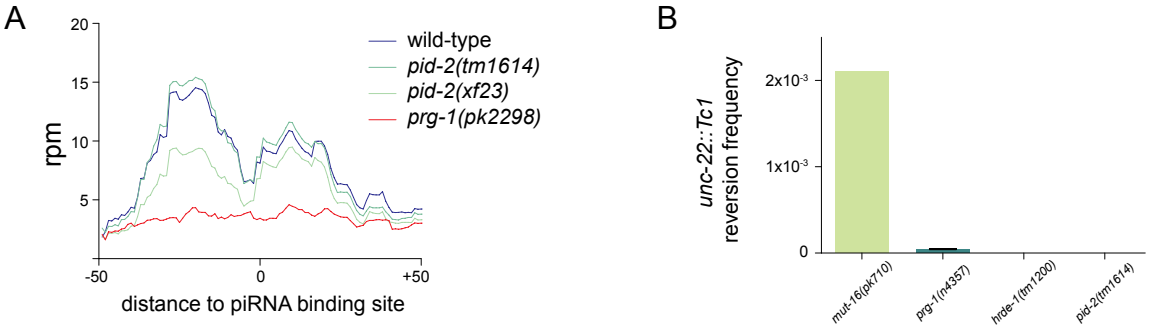


Figure 7. *pid-2* mutants show a PRG-1 dependent secondary 22G siRNA response, at endogenous PRG-1 recognition sites. Plot shows the average 22G coverage in a 100 bp window frame centered at the piRNA recognition site of WAGO-1 targeted genes. 22G siRNAs were considered as reads that start with a G, are 20 to 23 nucleotides long and mapped antisense to each piRNA target. piRNA targets were determined as described by Lee and colleagues (Lee et al., 2012). Both *pid-2* mutant alleles show secondary siRNA response triggered by PRG-1 recognition.

is no longer confined near the piRNA binding site and spreads through the transgene (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). In wild-type worms carrying the piRNA sensor in the RNAe state we see a local 22G siRNAs population near the *21ur-1* recognition site and a second 22G siRNAs population mapping the GFP portion of the transgene (figure 6A). As expected, in *prg-1* mutants there is a loss of 22G siRNAs near the piRNA recognition site (figure 6B). However, in *pid-2* mutants *prg-1* recognition is still able to trigger 22G siRNAs production (figure 6C, 6D and 6E).

pid-2 mutants have PRG-1 triggered secondary siRNA response

We then asked if this local 22G production in *pid-2* mutants is also present in endogenous *prg-1* targets. For that we searched for piRNA binding sites as described by Lee et al (Lee et al., 2012). From all identified targets we selected the ones that belong to WAGO-1 targeted genes and looked at the 22G siRNA levels in a 100 bp window centered in the potential piRNA binding sites (figure 7A). As previously described (Lee et al., 2012), wild-type worm show a 22G siRNA population near the piRNA binding site that is dependent on *prg-1*. Similar to what we observe with 22G siRNAs mapping the sensor at the *21ur-1* binding site, 22G siRNA levels near endogenous piRNA binding sites are unaffected in *pid-2* mutants. These observations indicate that the triggering of a mutator response by *prg-1* is seemingly normal in *pid-2* mutants.

Since *pid-2* mutants have local 22G siRNA production at both the sensor and endogenous piRNA biding sites, but fail to silence the piRNA sensor, we asked if *pid-2* mutants are able to silence *Tc1* transposons. In chapter 3 we showed that *Tc1* transposon is de-silenced when we simultaneously remove HRDE-1 and PRG-1 activity. In *hrde-1* mutants, PRG-1 activity is enough to silence *Tc1* activity as we saw no reversions of a strain carrying a *Tc1* insertion in the *unc-22* gene (figure 7B). However, *hrde-1* alone is not enough to keep *Tc1* silenced since *prg-1* mutants show a low level of *Tc1* mobilization (figure 7B). Likewise, we failed to see any *unc-22::Tc1* reversions in *pid-2* mutants. This is in line with the fact in *pid-2* mutants, PRG-1 still triggers secondary 22G siRNAs production, and further strengths the indication that these mutants still have a functional piRNA pathway.

FINAL REMARKS

We described the isolation of a novel factor involved in piRNA induced silencing: *pid-2*. *pid-2* mutants not only fail to fully de-silence the piRNA sensor but also show seemingly wild-type small RNA populations. A similar weak sensor silencing is observed in *henn-1* mutants (Luteijn et al., 2012); however these become silenced through RNAe after a few generations, while we have not seen any transition to RNAe in *pid-2* mutants carrying the sensor. The observed phenotypes raise two questions: firstly, why is the secondary siRNA response, triggered by *prg-1*, not enough to fully silence the sensor, and secondly, if there is still mutator activity, why is there no onset of RNAe at the piRNA sensor as observed in *henn-1* mutants.

The answer to the first question might come from the structure of the sensor itself. It has been shown that in worms, the Argonaute protein CSR-1 counteracts silencing activity (Claycomb et al., 2009; Seth et al., 2013; Wedeles et al., 2013). In the piRNA sensor, the *21ur-1* recognition site is just downstream the H2B coding sequence. This proximity to the H2B may allow CSR-1, up to certain point, to counteract *prg-1* silencing. This balance of silencing vs protection would allow partial sensor silencing and explain why *Tc1* activity is still fully repressed in *pid-2* mutants.

As for the second question, the answer might be the *pid-2* function. Taken the fact that in *pid-2* mutants both piRNA pathway and RNAe pathways are seemingly unaffected and that the sensor fails to “enter” RNAe, we propose a model where *pid-2* is bridging mutator activity triggered by *prg-1* recognition and the establishment of RNAe, but it is not essential for RNAe maintenance. If correct, this model ultimately implies that mutator genes act in two independent complexes: one producing “local” 22G siRNAs from primary pathway inputs such as PRG-1 or ERGO-1 targeting and the other producing 22G siRNAs from nuclear RNAi cues.

ACKNOWLEDGEMENTS

We thank Emil Karaulanov for help in bioinformatics analysis and Chung-Ting Han for help in generating small RNA libraries and sequencing. This work was supported by a Deutsche Forschungsgemeinschaft grant KE 1888/1-1 (Project Funding Programme; R.K.) and a grant from Fundação para a Ciência e Tecnologia (FCT)SFRH/BD/51001/2010 (B.A.).

MATERIALS AND METHODS:

Small RNA isolation

100uL of synchronized C. elegans pregnant adults were washed in M9 buffer (22nM Na2HPO4, 33 mM KH2PO4, 86 mM NaCl, 1 mM MgSO4) and digested in lysis buffer (200mM NaCl, 100 mM Tris pH 8.5, 50 mM EDTA, 0.5 % SDS, 200 ug/mL Prot-K) for 3h at 65°C followed by 15 min at 95° C to denature the Prot-K. Lysate was then incubated with DNase I (NEB) for 30 min at 37° C. Total RNA was then isolated using TRIZOL-LS using the manufacturer instructions. Samples where further enriched for the small RNA fraction using the miRvana miRNA isolation kit (life technologies).

Library preparation

2 ug of small RNA was treated with 5 U of tobacco acid phosphatase (Epicenter) at 37°C for 2 h to digest 5' tri- and di-phosphates to mono-phosphates. RNA was size-selected between 15- to 35-nt on 15% TBE-urea gel. Gel-purified RNA was eluted overnight in 300 mM NaCl and then precipitated with 100% isopropanol and Glycoblue for 1 h at -20°C. The pellet was washed once with 75% ethanol and dissolved in nuclease-free water. Then, this purified fraction was confirmed by Bioanalyzer Small RNA chip (Agilent). Library preparation was based on the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England BioLabs) with slight modification. In brief, small RNA was first ligated to the 3' adapter and then the 5' adapter, both of which contained 4 random bases and were chemically synthesized by Bioo Scientific. Adapter-ligated RNA was reverse-transcribed and PCR-amplified for 14 cycles using index primers. The PCR-amplified cDNA construct was purified using AMPure XP beads (Beckman Coulter). The purified PCR reaction was checked on the Bioanalyzer using High Sensitivity DNA chip (Agilent). Size selection of the small RNA library was done on LabChip XT instrument (Perkin Elmer) using DNA 300 assay kit. Only the fraction containing 140-165 bp was pooled in equal molar ratio. The resulting 10 nM pool was denatured and diluted to 10 pM with 5% PhiX spike-in and sequenced as single-read on HiSeq 2500 (Illumina) in rapid mode for 50 cycles using on-board cluster generation. After demultiplexing, on average 35 million passing filter reads were obtained per sample.

Data analysis

The raw reads in FastQ format were filtered from 3' adapter sequences and size-selected in the range 15-35 bases (plus 8 bases random barcodes) using cutadapt v.1.2.1 (Martin, 2011) using parameters: -a AGATCGGAAGAGCACACGTCT -O 8 -m 23 -M 43 . Subsequently, PCR clonal reads were deduplicated using Bash and Awk commands. All reads containing low-quality (Phred+33 score less than 20) bases were filtered with the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/; fastq_quality_filter -q 20 -p 100 -Q 33), then the files were reformatted from FastQ into tabular format, sorted and deduplicated based on full sequence identity (library insert plus 5' and 3' random barcodes of 4 nucleotides), and finally converted back to FastQ format for mapping. Quality assessment of the raw and processed data was done with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Mapping to the C. elegans genome reference WS244 was performed using Bowtie v1 (Langmead et al., 2009) with parameters: -v0 -M1 --best --strata --nomaqround --tryhard --trimm5 4 --trim3 4. All mapped reads where then annotated using bedtools intersect (Quinlan and Hall, 2010) with a customized WS244.gff3 (ftp://ftp.wormbase.org/pub/wormbase/releases/WS244/species/c_elegans/PRJNA13758/c_elegans/PRJNA13758.WS244.annotations.gff3.gz); using the following parameters: -abam -b[custom annotated gff3] -bed -wa -wb. Mapped and annotated reads were subsequently filtered for size and starting nucleotide using custom made python scripts and normalized to total of non-structural reads between 18 and 30 nucleotides. Structural reads were considered as reads that mapped rRNAs, tRNAs and snoRNAs. During the analysis we considered miRNA reads that were 22 to 24 nt long sense to annotated miRNAs, we considered piRNA reads that were 21 nt long, started with a "T" and were sense to annotated piRNAs, and 22G reads that were 20 to 23 nt long, started with a "G" and map antisense to genes, transposable elements or pseudogenes. In order to categorize genes, we retrieved WAGO-1 targets, ERGO-1 targets, mutator targets, ALG-3/4 targets and CSR-1 targets from (Claycomb et al., 2009; Conine et al., 2010; Gu et al., 2009; Phillips et al., 2014; Vasale et al., 2010) respectively. For each protein coding gene saw if it was a WAGO-1 target, if not we saw if it was an ERGO-1 target, if not we saw if it was a mutator target and so on, in the previously mention order. The validation of a gene in a category would automatically exclude it from the following categories; this was to insure that there would be no duplication, since some categories overlap

partially.

Conservation analysis

Each protein sequence was blasted against the protein database of each species using BLASTP 2.2.29+ with default settings. For each protein, only the best hit in each species was considered, hits with an e-value higher than -20 were discarded. For each selected hit, the bitscore was used as a to infer about conservation by normalizing the bitscore of each species to its C. elegans homologue. Futher information about the database used for each species can be seen in supplemental table 3.

RT-qPCR

Worm cDNA was synthetized from total RNA from synchronized adults using ProtoScript M-MuLV First Strand cDNA synthesis Kit (NEB) following the manufacture instructions for Oligo d(T) priming . GFP and TBB-2 amplicons were amplified in a Viia7 Real Time PCR System using the following primers:

GFP_forward: CTACCTGTTCCATGGCCAAC;

GFP_reverse: GGCATGGCACTCTTGAAAAA;

tbb-2_forward: CAATGTGCTTGACGTGATCC;

tbb-2_reverse: AGGTTGGGTTGGTGAGTTTG.

REFERENCES

De Albuquerque, B.F.M., Luteijn, M.J., Cordeiro Rodrigues, R.J., van Bergeijk, P., Waaijers, S., Kaaij, L.J.T., Klein, H., Boxem, M., and Ketting, R.F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28, 683–688.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150, 88–99.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M.E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E. a (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science (80-.). 337, 574–578.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31, 67–78.

Billi, A.C., Alessi, A.F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., and Kim, J.K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. PLoS Genet. 8, e1002617.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van Wolfswinkel, J.C., Chaves, D. a, Shirayama, M., Mitani, S., Ketting, R.F., et al. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139, 123–134.

Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 107, 3588–3593.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell 31, 79–90.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94–108.

Goh, W.S., Seah, J.W.E., Harrison, E.J., Chen, C., Hammell, C.M., and Hannon, G.J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. Genes Dev. 28, 797–807.

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J., et al. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell 36, 231–244.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151, 1488–1500.

Kamminga, L.M., van Wolfswinkel, J.C., Luteijn, M.J., Kaaij, L.J.T., Bagijn, M.P., Sapetschnig, A., Miska, E.A., Berezikov, E., and Ketting, R.F. (2012). Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans. PLoS Genet. 8, e1002702.

Kasper, D.M., Wang, G., Gardner, K.E., Johnstone, T.G., and Reinke, V. (2014). The C. elegans SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. Dev. Cell 31, 145–158.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150, 78–87.

Luteijn, M.J., and Ketting, R.F. (2013). PIWI-interacting RNAs: from generation to transgenerational epigenetics. Nat. Rev. Genet. 14, 523–534.

Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO

J. 31, 3422–3430.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell 136, 656–668.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10.

Montgomery, T.A., Rim, Y.-S.S., Zhang, C., Dowen, R.H., Phillips, C.M., Fischer, S.E.J., and Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. PLoS Genet. 8, e1002616.

Phillips, C.M., Montgomery, B.E., Breen, P.C., Roovers, E.F., Rim, Y.-S., Ohsumi, T.K., Newman, M.A., van Wolfswinkel, J.C., Ketting, R.F., Ruvkun, G., et al. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. Curr. Biol. 24, 839–844.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127, 1193–1207.

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Dev. Cell 27, 656–663.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). PiRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150, 65–77.

Spencer, W.C., Zeller, G., Watson, J.D., Henz, S.R., Watkins, K.L., McWhirter, R.D., Petersen, S., Sreedharan, V.T., Widmer, C., Jo, J., et al. (2011). A spatial and temporal map of C. elegans gene expression. Genome Res. 21, 325–341.

Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U. S. A. 107, 3582–3587.

Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. Curr. Biol. 18, 861–867.

Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013). Protection of Germline Gene Expression by C. elegans Argonaute CSR-1. Dev. Cell.

Weick, E.-M., and Miska, E. a (2014). piRNAs: from biogenesis to function. Development 141, 3458–3471.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R. a, Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E. a (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. 28, 783–796.

SUPLEMENTAL INFORMATION

Supplemental table 1. Non-synonymous mutations present in the strain carrying the *xf23* allele.

gene	a.a change	enriched in PGC	gene	a.a change	enriched in PGC
F58D5.7	A214T	no	C51F7.1	T319I	no
C01A2.4	L60F	no	W07G4.8	R31Q	no
Y71G12B.23	A62T	no	C56A3.7	P167S	no
Y47G6A.12	E1219K	no	C06B3.4	P286S	no
R12E2.8	G114S	no	T02B5.3	G462E	no
C10H11.9	S703N	no	ZC376.3	G567E	no
F28B3.1	H317Y	no	T26H5.3	G97E	no
F28B3.1	A327V	no	F36G9.13	A94T	no
F28B3.7	G113R	no	Y40B10A.6	G76R	no
C34G6.7	G425E	no	T28A11.7	G213E	no
K02F2.1	A920V	no	Y61A9LA.4	A140T	no
F20G4.1	V572I	no	T01G6.5	P78L	no
Y48G1C.1	W122*	yes	W06H8.8	Q7214K	no
T24D1.4	I31F	no	W06H8.8	T4611I	no
F38A3.1	G204R	no	W06H8.8	A4604D	no
F35D11.11	E1453K	no	F43H9.3	Q14*	no
ZK1248.1	P424L	no	F45F2.1	G452E	no
C27H5.3	P32Q	no	Y50D4C.4	C57Y	no
ZK1128.8	G99R	yes	C50C10.4	S189N	no
H04D03.1	D177E	no	F21A10.2	S553N	no
Y111B2A.24	A81V	no	R07E3.5	G262E	no
F59A2.6	G382E	no	C33D3.1	P206L	no
C14B1.10	P325L	no	F31B12.3	M329I	no
Y55B1AL.3	R207Q	no	F31B12.3	M329R	no
T12A2.15	G210E	no	R08B4.3	V135I	no
C56G2.3	A92V	no	T13G4.3	E300K	no
F37A4.4	A325V	no	R09A8.1	R492C	no
C18H2.1	K391E	no	C02B4.2	G117R	no
K02F3.6	S253N	no	M03B6.3	A191T	no
C38C10.5	P1455S	no	C31E10.1	A302T	no
B0464.2	T496I	no	C31E10.1	G235D	no
F01D4.3	E68K	no	F19G12.2	D487N	no
C29E6.1	E347*	no	C11H1.2	G320R	no
Y77E11A.12	S42F	yes	C44H4.7	P375L	no
T21D12.4	Y372N	no	T10B10.1	G111R	no
Y69A2AR.31	P667S	no	C33A11.1	P444S	no
Y54G2A.22	A292V	yes	F59F4.1	W30*	yes
Y67D8C.5	P530L	no	F41G4.4	G11R	no
Y67D8C.2	E92D	no	F52H2.1	V54I	no
W08E12.7	V318M	yes	C46C11.1	P163L	no

Y37E11AR.7	A237T	no	C46C11.1	Q714*	no
F33D4.4	G99D	no	C42D8.2	T369I	no
T20D3.8	G24E	no	T23F2.2	E96K	no
MTCE.35	S340K	no	C15H9.2	R98C	no
Y32F6A.3	K6N	no	C03B1.13	G9R	no
ZC302.1	L241F	no	F08F1.5	R113K	no
Y50D4B.4	N280K	no	F08F1.8	E138K	no
K12F2.1	G647E	no	C03G5.7	P5L	no

The isolated strain carrying the *xf23* allele was de novo-sequenced. The table indicates the non-synonymous mutations. The third and fourth columns indicates if the mentioned gene is enriched in PGC as published by spencer and colleagues (Spencer et al., 2011).

Supplemental table 2. General statistics of small RNA sequencing libraries.

genotype	sample	total	mapped	non structural
21U sensor	1	8485740	6689943	2854746
21U sensor	2	7598752	6100628	1673176
21U sensor	3	6934542	5336487	1577523
21U sensor	4	7621136	5809665	1559442
prg-1(pk2298); 21U sensor	5	8984187	6868003	2107091
prg-1(pk2298); 21U sensor	6	7569818	5515688	1501608
pid-2(tm1614); 21U sensor	7	9743408	8059164	3030668
pid-2(tm1614); 21U sensor	8	5751018	4575956	1284766
pid-2(tm1614); 21U sensor [RNAe]	9	6380068	5075030	2070415
pid-2(tm1614); 21U sensor [RNAe]	10	5866897	4570882	1695395
pid-2(xf23); 21U sensor	11	6867325	5497598	1602658
pid-2(xf23); 21U sensor	12	7808168	6414677	2942696

List of articles

De Albuquerque, B.F.M., Luteijn, M.J., Cordeiro Rodrigues, R.J., van Bergeijk, P., Waaijers, S., Kaaij, L.J.T., Klein, H., Boxem, M., and Ketting, R.F. (2014). **PID-1 is a novel factor that operates during 21U-RNA biogenesis in *Caenorhabditis elegans*.** Genes Dev. 28, 683–688

De Albuquerque, B.F.M., Placentino, M., and Ketting, R.F. (2015). **Maternal piRNAs are essential for germline development following de-novo establishment of endo-siRNAs in *Caenorhabditis elegans*.** Dev. Cell *[in revision]*